University of Salford MANCHESTER

MSc by Research in Environmental Studies

Signatures of Selection in the Red Fox (*Vulpes vulpes*) in Europe using genotype-by-sequencing

Liam Roberts

2019

School of Environment & Life Sciences

MSc by Research Thesis

Contents

List of Tables

List of Figures

Figure 7 - The populations Fin_S, Swe_NE, Swe_S, and Nor_NE show a distinct change in the allele frequency at position 13229819 on chromosome 27 from a guanine (G) to a cytosine (C). The remaining 26 fox populations have a guanine (G) fixation (1) allele frequency.

List of Appendices

Appendix I - Outliers that tested for positive diversifying selection and balancing selection using the BayeScan method.

Appendix II - Candidate genes that was found to be uncharacterised.

Appendix III - Details of each candidate gene found including SNPs, full gene name, reference and alternate allele (major and minor allele), E-value from BLAST software, whether the SNP occurred in exon or intron, FST, alpha-values, and other information from BayeScan, P-value from PCAdapt, and strand of nucleotides surrounding target SNP from SAM file.

Appendix IV - Results from candidate gene list used in g:Profiler with enrichment test showing terms with p-values of <0.05 after correction with multiple testing using the Benjamini-Hochberg FDR method.

Appendix V - Results from Ease score of >1.3 (P-value <0.05) the Functional Annotational Clustering candidate gene list used in DAVID.

Appendix VI - Candidate genes list in each term from DAVID (see Appendix V).

Appendix VII - References for each gene summary in Appendix IX.

Appendix VIII - Allele frequency from each of the 161 SNPs used in this study.

Appendix IX - Candidate gene found in this study. A phenotypic trait shown to be influenced by each one of these genes by a previous study along with the system/organism used to gather this information. Finally, the FST for each SNP found in this study for each particular gene shown is also displayed. For more detail see Appendix III, IV, V, VI, VII, and VIII.

Acknowledgements

I would like to thank my supervisor Dr Allan McDevitt for his advice and helping me shape this project, as well as his understanding and support. I would also like to thank Sam Browett and Dr Ilaria Coscia for their initial help in figuring out Plink and other computer support which helped kick start my focus in bioinformatics.

Statement of Originality

I declare that, with the exception of any statements to the contrary, the contents of this report/dissertation are my own work, that the data presented herein has been obtained by experimentation and that no part of the report has been copied from previous reports/dissertations, books, manuscripts, research papers or the internet.

Signed.....

Print name... Liam Roberts .....

Date.27/06/2019

Glossary of Terms

BayeScan – a statistical method to identify loci outliers and form of selection (diversifying or balancing selection) based on the multinomial Dirichlet model.

Candidate gene- a gene or a variation of a gene which may be involved with a phenotypic trait of interest.

FST – fixation index that measures the differentiation between and within populations based on genetic structure.

GBS - genotyping-by-sequencing is a method which is used to identify single nucleotide polymorphisms (SNPs) in the genome.

LD - linkage disequilibrium is used to identify the non-random association of alleles at different loci in a specific population.

PCA – displays genetic distances between individuals or populations.

PCAdapt – a statistical method used to identify genetic markers involved in biological adaptation.

SNP/s - Single nucleotide polymorphism/s are the most common type of genetic variation in which a difference in a single DNA nucleotide is found.

Abstract

Evolution acts on variations in individuals via the mechanisms of natural selection, genetic drift, and gene flow. Using genomic sequencing methods such as genotype-by-sequencing (GBS) to detect single nucleotide polymorphisms (SNPs) can lead to a better understanding of a species evolutionary history and identify patterns of local adaptations. Abiotic and biotic factors can cause selective pressures that may be seen in the red fox's genome. Additionally, debate continues to whether red foxes remained in isolated refugia or existed as a continuous population in southern Europe. Regardless, previous studies highlight the importance of refugia of the Last Glacial Maximum period in shaping genetic variation and differentiation in many taxa. This study aims to provide additional evidence to this dispute while also uncovering signatures of selection in genes related to the new conditions faced as red foxes pushed northwards in a warming climate. To test these predictions, recently generated GBS data from over 500 red foxes (*Vulpes vulpes*) in Eurasia was utilized. First, I formed a population criterion based on the sample's location and the fox's ecological habits. This resulted in 436 samples in which I identified 30,708 SNPs after bioinformatic filtering. I then ascertained population structure using principal component analysis (PCA). This provided consistencies with red foxes living outside of proposed refugia's during the Pleistocene. To detect candidate genes for diversifying selection I used outlier loci that overlapped from both PCAdapt and BayeScan methods. This revealed candidate genes relating to immunity, vision, and other phenotypic traits. A particularly interesting result found a significant allele frequency change in just four Nordic fox populations in the innate immune MUC19 gene which may have responded to a parasitic selective force. Further investigation into these findings may elucidate clearer, more detailed adaptive processes in the wild red fox within Eurasia.

Keywords: Red fox; Selection; Adaptation; PCAdapt; BayeScan; PCA; Allele Frequency

1. Introduction

1.1 Statement of the Problem

There seems to be a complete lack of genome-wide studies (at the time of writing) in the red fox (*Vulpes vulpes*) across the European (Eurasian) continent. This study aims to address this situation by taking previously generated GBS data (provided by A. McDevitt, *unpublished data*) and using bioinformatic filtering and two different outlier methods to produce a new data set. This new data set was used to search for signatures of selection in the red fox. A previous genome-wide study also using SNP data on another wide ranging species of canid (Schweizer, et al., 2016) has found candidate genes relating to many phenotypic traits in different populations including metabolism and morphology. Therefore, it was expected that candidate genes relating to local adaptations would also be found in red foxes.

In addition to using this data set to search for signals of selection in the red fox, additional evidence will be obtained regarding the red fox's phylogeographic history by using a Principle component Analysis (PCA). There are studies which have different views to whether the red fox was confined to traditional well-defined and separated refugia during the LGM and to what degree of population structure exists today due to their past. The discussion is likely to be debated far into the future. The results from this current study however are expected to provide a clearer picture of the fox's recent phylogeographic history.

By using genetic data to investigate the red fox's past we may uncover events or factors that may have shaped the evolutionary adaptations seen in today's foxes.

1.2 Evolution

Evolution takes place most often through mechanisms such as genetic drift, gene flow, and selection. Furthermore, biological evolution takes place at the population level as mixing of different combinations of genetic changes (alleles) oscillate in frequencies. Evolution then is at its basis is a change in allele frequencies in populations from one generation to the next (Johnson & Munshi-South, 2017).

Variations in the genome can lead to changes in the genotype which underlie changes in the phenotype and this can then affect the fitness of an individual (Ketterson et al., 2001). It is differences in fitness that allows natural selection to take place which in turn produces adaptations to changing environmental conditions of both abiotic (temperature, rainfall etc.) and biotic (predators, competitors, parasites etc.) factors. Fitness in evolutionary biological terms is broadly speaking an organism's (rarely populations or species) ability to survive and reproduce in their current environment (Orr, 2009).

1.3 Searching for Evolutionary Change

By focusing on genomic regions which seem to have been impacted by positive selection, genes important to individuals from a functional aspect can theoretically be identified in populations such as migratory behaviour in monarch butterflies (*Danaus plexippus*) (Zhan et al., 2014) or hypoxia adaptations in wolves (*Canis lupus chanco*) (Zhang et al., 2014). Unfortunately, when searching for genetic variation of interest the problem of high false positive rates may occur as closely related populations often share gene flow and population histories. This is often reflected in allele frequency correlations between the populations most closely related (Coop, Witonsky, Di Rienzo, & Pritchard, 2010). To partially alleviate the

problem of false positives due to gene flow or genetic drift multiple unlinked loci between populations can be compared. Schweizer *et al.*, (2016) took this approach in a study on the gray wolf (*Canis lupus*) in North America which found diversifying selection occurring despite gene flow also happening between populations. The technique of comparing unlinked loci between populations can be successfully used as selection usually occurs at locus specific sites while gene flow and genetic drift act across the whole genome (Nielsen, 2005). Particular outlying loci can then be statistically identified, and these can be presumed to be in Linkage Disequilibrium (LD) with genes, loci, or other genomic features that are under selection (Charlesworth and Charlesworth, 2017; Schweizer, *et al.*, 2016)

Linkage disequilibrium (LD) between unlinked loci can be caused by strong positive selection as can be seen in the lactase gene (LCT) within the human genome (Bersaglieri *et al.*, 2004; Poulter *et al.*, 2003). In particular, local adaptation together with strong selection may show LD between SNPs that have previously been associated with environmental variables such as precipitation, as presented by Long *et al.*, (2013) in the plant genome of *Arabidopsis thaliana*. Population genetic studies have found selection acting on subtle (metabolism-related genes etc.) and more obvious (growth related genes etc) phenotypes using this method. For example, Zhang *et al.*, (2014) focuses on positively selected hypoxia-related genes in wolves living in the high altitude (>3000 m) enviroments of the Qinghai-Tibet Plateau. From a post filtered data set of 266,299 SNPs, 84 hypoxia-related genes seemed to be under selection of which EPAS1, RYR2, and ANGPT1 were homozygous in the reference allele in wolves living in the lowlands and homozygous in the alternative allele in the highland (Tibetan) wolves. This, along with a high degree of LD surrounding these three genes suggests strong selection has taken place.

Less subtle phenotypic traits affected by selection are also known. For example, in some circumstances natural selection can be expected to act on increasing the size of members of a

species living in cold environments in relation to conspecifics inhabiting warmer environments. This is known as Bergman's rule (also see section 1.7 Ecology) and it seems that significant correlations between body mass and temperature in human populations do exist (Foster and Collard, 2013; Katzmarzyk and Leonard, 1998). Adaptations also seem to have arisen in populations of humans living at high altitudes which have genetic differences when compared with lowland human populations (Alkorta-Aranburu *et al.*, 2012; Eichstaedt *et al.*, 2014; Jeong *et al.*, 2014; Zhou *et al.*, 2013). Evidence for positive selection occurring in genomic regions of the human genome in regards to changes in diet, infectious disease, and other selective pressures can also be found (Akey, 2009; Sabeti *et al.*, 2006).

Genomic regions of particular interest are those in which the allele frequency variation is related to some external measurable difference among populations (Nielsen *et al.*, 2007; Schweizer et al., 2016). For example, Oakeshott, Wilson and Knibb, (1998) finds latitudinal associations in allele frequencies with Adh, Gpdh, and Est6 enzyme polymorphisms in *Drosophila melanogaster* (Oakeshott et al., 1998). Other studies with *D. melanogaster* show correlations to climatic variables in variants involved with phenotypes including energy metabolism (Oakeshott, Wilson and Knibb, 1998; Sezgin *et al.*, 2004) and longevity (Hudson *et al.*, 1994; Schmidt, Duvernell and Eanes, 2002) while variants involved with metabolic processes associated with high altitude are also known (Jha et al., 2016).

While studies in animals are becoming more common there are also many population genetic studies that have been done in plants. Fischer *et al.*, (2013) identified 175 genes which strongly associated with five (precipitation, slope, solar radiation, site water balance, and temperature) abiotic environmental factors in the model organism *Arabidopsis*. In the non-model Loblolly Pine (*Pinus taeda L.*) five loci in genes were identified associating with aridity gradients. The primary functions of these genes translated products were abiotic and biotic stress responses (Eckert et al., 2010).

Model organisms have long been widely used in population genetic studies in laboratory-based studies and more recently in natural population studies (Ekblom & Galindo, 2011). With more ecologically important model species genomes being sequenced thanks to advances in technology (next-generation sequencing) species such as sticklebacks (*Gasterosteus*) (Hohenlohe et al., 2010) water fleas (*Daphnia*) (Eads, Andrews, & Colbourne, 2008), zebrafish (*Danio rerio*) (Howe et al., 2013), the domestic dog (*Canis lups familiaris*) (Lindblad-Toh et al., 2005) and fruit flies (*Drosophila*) (dos Santos et al., 2015) can be used to map wild species genomes in population genetic studies. These species provide reference genomes to allow closely related wild species in natural populations to be studied allowing important questions to be asked related to conservation and ecology (Davey *et al.*, 2011; Wheat, 2010).

By using the dog genome as a reference genome Schweizer, *et al.*, (2016) was able to study the genome of the grey wolf. The study found signals of selection in genes relating to morphology, coat coloration and metabolism. Variation in allele frequencies relating to environmental changes in temperature and precipitation was also found and Schweizer, *et al.*, (2016) suggested this is more consistent with local adaptation rather than genetic drift. Other studies have also used a reference genome for sequence alignments to search for signals of selection. For example, the domestic sheep genome was used in wild big horn sheep which found evidence of a selective sweep for the RXFP2 gene which influences horn morphology and size in wild bighorn and Soay sheep respectively (Kardos et al., 2015). Another example with fish is the Atlantic salmon (*Salmo salar*) genome was used for assessing conservation issues with the brook trout (*Salvelinus fontinalis*) (Buonaccorsi, Malloy, Peterson, Brubaker, & Grant, 2017).

Population genetic studies attempting to answer questions on local adaptation begin by looking for signals of selection in a species genome by using genome-wide selection scans

(Oleksyk, Smith, & O'Brien, 2010). Restriction site-associated DNA sequencing (RADseq) is a group of next-generation sequencing methods that enables hundreds or thousands of polymorphic genetic markers to be studied in a single experiment covering the whole genome. The commonly used RADseq methods can be grouped into 2 main groups, one group (original RADseq method) sequences fragments adjacent to single restriction enzyme cut sites while the other group (more modern RADseq methods) sequences fragments flanked by two restriction enzyme cut sites. The most commonly used RADseq methods and their trade-offs are shown in Table 1 (Andrews et al., 2016). RADseq methods are invaluable in SNP discovery and genotyping in ecological and evolutionary genomic studies (Anderson, Willis and Mitchell-olds, 2012; Rellstab *et al.*, 2015; Savolainen, Lascoux and Merilä, 2013) however they can produce genotyping error and bias (Arnold, Corbett-Detig, Hartl, & Bomblies, 2013). When used effectively it can be a vital tool to study natural population genomics and produce deeper understandings in evolutionary, conservation, and ecological related questions (Andrews *et al.*, 2016; Fitzpatrick, Keller and Lotterhos, 2018). However, choosing an appropriate RADseq method based on the study system, budget, and specific research question is important as protocols differ (Andrews et al., 2016).

Radseq methods are not the only way to produce population genetic studies as the successful study of grey wolves over North America using a SNP array method has been done (Schweizer *et al.*, 2016). SNP arrays have been widely used for Genome Wide Association Studies (GWAS) and produce high quality greater marker densities which makes the results easier to analyse. However, SNP arrays suffer from ascertainment bias, prior genetic knowledge is needed, and the cost of sequencing is also much higher than in Genotyping by sequencing (GBS) or ddRAD for example (Scheben, Batley & Edwards, 2017). This makes SNP arrays less practical to keep up with the increase in population genetic studies being produced (Ekblom & Galindo, 2011).

Table 1. GBS and common RADseq methods and their trade-offs. Figure adapted from Andrews et al., (2016).

	GBS	Original RAD	2bRAD	ddRAD	ezRAD
Ability to modify number of loci	Change the restriction enzyme	Change the restriction enzyme	Change the restriction enzyme	Select the size of window or change the restriction enzyme	Select the size of window or change the restriction enzyme
Loci number per each 1 Mb of genome size	5-40	30-500	50-1000	0.3-200	10-800
Length of loci	≤300 bp	≤300 bp unless building with contigs then up to 1kb	33-36bp	≤300 bp	≤300 bp
Cost per indexed or barcoded sample	Small	Small	Small	Small	Large
Effort per indexed or barcoded sample	Little	Intermediate	Little	Little	Large
Proprietary kit usage	No	No	No	No	Yes
Identification of PCR duplicates	With paired end sequencing	With degenerate barcodes	No	With degenerate barcodes	No
Specific equipment needed	None	Sonicator	None	Pippin Prep	Pippin Prep
Suitability for complex or large genomes	Moderate	Good	Poor	Good	Good
Suitability for de novo locus identification (no reference genome)	Moderate	Good	Poor	Moderate	Moderate
Available from commercial companies	Yes	Yes	No	Yes	No

Whichever method is used, being able to identify genes involved in adaptive evolution enhances our comprehension of processes that drive evolutionary and phenotypic specialization while also uncovering evolutionary history aspects of a species or a group of species (Aguileta *et al.*, 2010; Anderson, Willis and Mitchell-olds, 2012; Kardos *et al.*, 2015). Therefore, using a RADseq method is a good place to begin gathering evidence of local adaptions and determine what functional variants are under selection in natural populations.

Genotyping by sequencing (GBS) is often used because it is a low cost widely useful and effective technique that has also been utilised in other population genetic studies (Davey et al., 2011; De Donato *et al.*, 2013; Huang *et al.*, 2014; Li *et al.*, 2014). Although GBS can be used in many circumstances this method can suffer from high levels of missing data (Scheben et al., 2017). However, unlike the intensive labour effort needed to produce a library preparation using original RADseq (Scheben et al., 2017) the GBS method is relatively simple to implement and the informatics pipelines are publicly accessible (De Donato et al., 2013).

It is true that the ddRAD achieves a greater coverage of a subset of the genome and so can offer a more effective SNP genotyping result than GBS or original RADseq. This is attained by using two restriction enzymes, (after size selection) recovering a specific subset of the genome and reducing the size of the subset sampled (Scheben et al., 2017). However, the GBS method can compensate for a lower, more uneven coverage, if provided with prior genomic information such as a reference genome. This compensation is achieved by using imputation (Scheben et al., 2017).

To use the GBS method samples must be taken which for example, can come from green tissue from plants (Annicchiarico et al., 2017) or blood from mammals or fish (De Donato et al., 2013: Li et al., 2014). The DNA collected from these samples can be taken and the GBS

method can be applied (below). GBS, like other RADseq methods, can be used for genetic mapping and genetic linkage which can help identify genetic loci leading to uncovering traits of interest (Scheben et al., 2017).

GBS works by first using a restriction enzyme to facilitate the digestion of large sequences of DNA into more manageable fragments just like the original RADSeq method (Andrews et al., 2016). GBS then uses barcoded adapters and common adapters (oligonucleotides) which allows the ligation of the DNA fragments, unlike many RADSeq methods GBS does not use "Y" or "forked" adapters for ligation (Elshire et al., 2011). GBS uses an indirect size selection unlike most RADSeq methods which directly select fragment size by using manual or automated gel cutting techniques or magnetic beads (Andrews et al., 2016). Primers are added with complementary sequences to amplify the sequences which may produce overlap with GBS as some fragments may be small. Then Polymerase Chain Reaction (PCR) is used to amplify the pool and overlaps increasing genotyping accuracy closer to the ends of the reads, where rates of sequencing error are highest (Andrews et al., 2016). This is followed by bioinformatic filtering and information on the genome's variation can then be gathered from populations composed of many individuals (Perea et al., 2016). Once the data has been mapped a sliding window analysis can (optionally) then be used to detect genetic regions of interest, notably regions that seem to be under divergent selection between populations (Andrews *et al.*, 2016; Zhan *et al.*, 2014; Zhang *et al.*, 2014;). Genes under selection can then be looked at in a broad sense and features of these genes can be determined using blast software (She, Chu, Wang, Pei, & Chen, 2008).

By using GBS together with outlier methods, SNPs can be identified that are presumed to be under selection whether the phenotypes are known or not (Chase *et al.*, 2009). The GBS method is used in plant breeding to improve complex traits such as yield, height, and drought tolerance (Annicchiarico *et al.*, 2017; Huang *et al.*, 2014; Poland and Rife, 2012; Pootakham

et al., 2015). While GBS has also been used in animal studies detecting traits such as sexual selective traits in barn swallows (*Hirundo rustica*) (Wilkins et al., 2016), responsive traits to infection in Aleutian mink (*Carnivore amdoparvovirus*) (Farid, Gardner, Butler, Rupasinghe, & Myles, 2014), and dairy traits (milk production etc.) and parasite resistance/susceptibility traits in cattle (De Donato et al., 2013).

The GBS method is often implemented along with FST outlier methods which are based on comparing FST across loci and against a demographic model then identifying those FST coefficients that are significantly different (Excoffier, Hofer, & Foll, 2009). In population genetics the composition of genetic differences between and within populations can be summarized by the fixation index (FST) (Weir & Cockerham, 1984). FST estimates are used for Single nucleotide polymorphisms (SNP) or microsatellites for wild populations in population genetic studies (Coates *et al.*, 2009; Renaut *et al.*, 2010; Schweizer *et al.*, 2016) . SNPs examined on a large scale can be used to find important variant changes by measuring FST patterns throughout a genome of a particular species (Zhan et al., 2014) or the genome of a closely related species such as in the case of the domestic dog (*C. l. familiaris*) and the grey wolf (*C. lupus*) (Schweizer, et al., 2016). Loci with high FST results above a neutral distribution value are presumed to be outliers and candidate genes for diversifying selection. However, the methods used to detect outliers differ in some ways with each having its own strengths and weaknesses (Narum and Hess, 2011). The method to be used depends on the question that is being asked and methods can be combined to strengthen the results obtained (Nielsen et al., 2007; Narum & Hess, 2011).

1.4 Population Structure and Natural Selection Assessment

Population genetic studies often use the program PCAdapt (Arnyasi *et al.*, 2017; Duforet-Frebourg *et al.*, 2016; Grant *et al.*, 2017) which uses a Principle component Analysis (PCA) along with SNP data as a tool to investigate geographic genetic variation (Patterson, Price and Reich, 2006; Pimsler, Jackson and Lozier, 2017; Zhang et al., 2014). PCA uses a multivariate analysis in which to ascertain population structure (Duforet-Frebourg *et al.*, 2016; Patterson, Price and Reich, 2006). PCA takes the data set of genetic variation and reduces it down into a number of principle components (*K* axes) and each principle component can represent an evolutionary processes such as divergence between populations (Duforet-Frebourg *et al.*, 2016; McVean, 2009). To obtain a number of principle components (*K*) a scree plot is used (Holand, 2016). It is important to correctly group individuals into populations when looking for signals of selection as it reduces the likelihood of missing these signals (Duforet-Frebourg et al., 2016). The PCAdapt method uses a PCA to produce a hierarchical population structure and also uses the Mahalanobis distance test statistic (McLachlan, 1999) which increases the power to give more reliable results from a genome scan (Luu et al., 2017). This test statistic works by taking the resulting *K* z-scores to create a vector, this then measures how closely related a SNP is to the first K principal components. Finally the Mahalanobis distance is calculated for each SNP to detect outliers (vector of z-scores that deviate from the distribution of the main bulk of points) (Luu et al., 2017).

To strengthen the evidence for diversifying selection a Bayesian statistical analysis in the program BayeScan which implements the Markov Chain Monte Carlo algorithm method described by Foll and Gaggiotti, (2008) can also be used to detect candidate loci. Additionally, these candidate loci can be differentiated into those that appear to have occurred due to diversifying selection from those loci that appear to be due to balancing

selection (Foll & Gaggiotti, 2008). This is estimated by using the ratio of the posterior probability of two models (selection/neutral) based on the data and is used in many population genetic studies looking for signals of selection (Fischer *et al.*, 2014; Freedman *et al.*, 2010; Poulsen *et al.*, 2011; Schweizer *et al.*, 2016) . Positive or negative alpha values are produced which implies diversifying selection or balancing selection has taken place respectively and for each locus an FST estimate is plotted against the P-value. FST results are related to the strength of selection and high FST values are associated with diversifying selection in geographically restricted areas while balancing selection is more consistent with a low FST (Beaumont and Balding, 2004; Ronald and Akey, 2005). This method of distinguishing positive diversifying selection from balancing selection is consistent with studies involving wild natural populations (Alberto *et al.*, 2013; Fischer *et al.*, 2011; Guo, Li and Merilä, 2016; Schweizer *et al.*, 2016).

Natural selection is expected to leave genes with a signature pattern of neutral polymorphisms in neighbouring genomic regions. The pattern is expected to show a reduction in the genetic diversity among populations and increase in genetic diversity within populations in balancing selection while with diversifying selection the reverse is expected (Charlesworth, Nordborg and Charlesworth, 1997; Nielsen, 2005). Studies seem to show that balancing selection is important in a few key classes of genes mostly relating to parasite-host interactions, recognition of kin, self-incompatibility in plants, and mating-type genes in fungi (Aguilar *et al.*, 2004; Andres *et al.*, 2009; Asthana, Schmidt and Sunyaev, 2005; Delph and Kelly, 2014; 2014; Devier *et al.*, 2009; Fijarczyk & Babik, 2015; Spurgin and Richardson, 2010). Furthermore, balancing selection may be more ubiquitous than has been previously supposed or shown. However, there are currently many difficulties in using polymorphism and divergence data analysis to figure out how widespread or important the true role of balancing selection really is (Fijarczyk & Babik, 2015).

Difficulties arise in the reliability in distinguishing loci under balancing selection in outlier methods, FST methods generally have comparatively little power in detecting balancing selection compared to detecting diversifying selection (Beaumont and Balding, 2004; Ronald and Akey, 2005). High false positive rates are also associated with detecting balancing selection (Excoffier, Hofer and Foll, 2009; Fijarczyk and Babik, 2015) and BayeScan in particular can experience excessive false-positive rates compared to when detecting diversifying selection (Lotterhos & Whitlock, 2014).

The relative advantages in detecting diversifying selection over balancing selection using outlier FST methods has led to successful non-model studies in natural populations to focus on diversifying selection only (Kardos *et al.*, 2015; Pilot *et al.*, 2014; Schweizer *et al.*, 2016). Thus, by concentrating on the loci presumed to have been selected only by diversifying selection, different aspects in the environment can be looked at in relation to changes in the genome. Changes in the biotic and abiotic environment may cause a shift in the allele frequencies to respond. By responding to environmental pressures, evolution by natural selection can produce populations which contain a large proportion of individuals with a phenotype (and genotype) that maximises fitness in their local environment (Einum & Fleming, 2000).

As evolution occurs over generations it is important to understand the life history of the species being studied. By looking at the distribution, ranging habits, and the ecology of the red fox we can begin to build a picture of what may have influenced the evolutionary course of this species in more detail. The red fox is a good candidate to look for local adaptation as they occur in a wide range of habitats across the European continent (Statham, Edwards, Norén, Soulsbury, & Sacks, 2018). This means different groups of foxes over their range will have come in to contact with different pathogens, prey and terrain and so be subject to different selective pressures.

1.5 Origin and Distribution of the Red Fox

The red fox (*V. vulpes*) belongs to the Canidae family as does the gray wolf (*C. lupus*), they both share a most recent common ancestor estimated at 10-12 million years ago (MYA) (Wayne, 1993; Williams, Muñoz-Garcia, Ostrowski, & Tieleman, 2004).

More recently fossil records suggest that the red fox was present in Europe at least 300,000 years ago during the Late Pleistocene and evidence of red fox remains indicate they possibly inhabited Europe over 400,000 years ago (Kutschera et al., 2013; Statham et al., 2014). Mitochondrial sequence DNA also points to red foxes initially diversifying in Eurasia (Holarctic lineage) before colonizing North America and Japan (Kutschera et al., 2013).

The extraordinary adaptability of the red fox enables it to inhabit a broad range of habitats, this is evident from relatively common partially fossilized bone remains (subfossils) from the Pleniglacial time (72–13 thousand years ago) found in locations such as, Spain, England, France, Italy, and Germany (Huijzer and Vandenberghe, 1998; Sommer & Benecke, 2005).

The Pleistocene was a time of a rapidly fluctuating climate which would have influenced the need to adapt to environmental changes for the wildlife (Barker, 2005; Huijzer & Vandenberghe, 1998). It is thought that intraspecific genetic variation has been strongly influenced during the cold periods of the ice age in Europe (Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998). Climatic extremes forced many plants and animals to retreat to areas where the environment was somewhat more habitable, these areas are known as refugia (Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998). During these periods, populations moved between being fragmented and restricted from one another to partially re-merging with one another (Statham et al., 2014).

The Pleistocene ice sheets reached their maximum ~27-19,000 years ago known as the Last Glacial Maximum (LGM). Postglacial recolonizations from glacial refugia at this time would most likely have major consequences for the evolution and ecology of today (Nadachowski & Sommer, 2006). Isolated populations will be subject to different selective pressures and so it is important to identify the locations of refugia during this period. Molecular genetic evidence from mammal, amphibian, arthropod, and plant taxa point to three model Mediterranean refugia existing during the Pleistocene which are the Iberian, Balkans, and Italian peninsulas (Hewitt, 1999; Nadachowski & Sommer, 2006; Taberlet et al., 1998). Additionally, two more refugia, the Carpathians mountains region and Dordogne region in France, have been identified (Herman *et al.*, 2017; McDevitt *et al.*, 2012).

As refugia populations were isolated from one another local adaptations and population differentiation had the opportunity to occur. With a changing climate gene flow had the chance to occur (Nadachowski & Sommer, 2006). As time passed in these independent populations before being reunited clear genetic differences should be inferred (signatures of selection) which are characteristic of a particular group. The red fox is a species that has a current broad range and so it would be expected that distinct evolutionary histories and a diverse population structure would exist (in the genome) if they were isolated in proposed refugia in the past (Kutschera et al., 2013).

However, Statham et al., (2018) in concordance with Edwards et al., (2012) points to red foxes not being restricted to glacial refugia during the LGM, though distinct patterns of colonization can be observed in the genome. The recent study by Statham et al., (2018) took 288 continent wide samples from European red foxes and used nuclear microsatellite loci and mitochondrial DNA to describe the continent's population structure. They found that the Iberian red foxes derived from European populations before the last global glacial maximum and contributing very little to other western European populations gene pool. This along with

mtDNA haplotypes estimates has indicated the Iberian Peninsula has been isolated since the last glacial maximum in the Pyrenees of ca. 50–70 Kya. Statham et al., (2018) also found the red foxes from the Italian peninsula were genetically distinct and isolated from the Iberian Peninsula while also contributing to the gene pool of neighbouring populations in central Europe and the Balkans. British and Irish red foxes were suggested to have been founded by populations from France (and the Netherlands) before the ice and land bridges from the continent disappeared under water. The rising water also isolated these British and Irish fox populations from one another. The Scandinavian Peninsula seems to have been colonised both from central European populations as they moved north via Denmark and from the east before becoming isolated from the central European populations due to sea level rise. Red foxes have shared the European continent with humans since the Pleistocene (Falguères, 2003) and it seems likely that foxes avoided humans as humans hunted them in the past as they continue to do so in some areas (Vuorisalo, Talvitie, Kauhala, Bläuer, & Lahtinen, 2014).

A recent change in human attitudes towards foxes (Plumer, Davison, Saarma, & Cameron, 2014; Scott et al., 2014) has allowed the existence of red foxes to live closer to human settlements. Urbanization of red foxes seems to be an evolutionarily recent event, moving into cities in southern England within the last 100 years then subsequently either independently occurring in other cities within its geographic range or spreading from this area (Scott *et al.*, 2014; Vuorisalo et al., 2014). These ‘urban’ foxes have been scientifically studied since the 1930s in Britain (Harris & Rayner, 1986).

With urbanization, red foxes became exposed to pollutants. Pollution resulting from urban development has many negative effects on wildlife ranging from survival rates, immune health, impaired reproduction, and other physiological stresses (Ditchkoff, Saalfeld, & Gibson, 2006). Endocrine disrupting chemicals (EDCs) and heavy metals also have damaging

effects (Corsolini *et al.*, 1995; Dip *et al.*, 2001; Dip et al., 2003; Pascoe, Zodrow and Greutert, 2014). The effects of these pollutants range from neurological damage, birth defects, cancer, immune suppression, cognitive and motor learning abilities, neurobehavioral and cognitive deficits in humans and other animals (Annamalai & Namasivayam, 2015; Ditchkoff et al., 2006; Erren, Zeuß, Steffany, & Meyer-Rochow, 2009; Dip et al., 2001) to lowered milk production and a decrease in survival of red fox kits (Pascoe et al., 2014).

Despite the challenges faced, red foxes are adaptable having increased their range over time and now they are native to four of the world's continents including in many cities of, Europe, Asia, North America, North Africa, and have been introduced into the continent of Australia where they now also thrive (Statham et al., 2014).

1.6 Home Ranges and Dispersal of Red Foxes

One aspect that enabled red foxes to expand their range at the end the LGM was the ability to disperse large distances. Dispersal rates vary in red fox populations which may be as far as 140km in Denmark (Harris, Trewhella, & Harris, 2015) and in Sweden up to 250km dispersal rates have been documented (Holmala & Kauhala, 2006). Goszczyński, Misiorowska and Juszko, (2008) also mentions young fox's dispersing in excess of several dozen kilometres. It is mainly juveniles or subadult red foxes that disperse (Goszczyński, Misiorowska, & Juszko, 2008) and Zimen, (1984) mentions a study in Germany and France that has shown they may travel as much as 10-100km in one to several nights. Furthermore, a study by Walton et al., (2018) documented the longest dispersal distance of any recorded red fox at a cumulative distance of 1036km during a 100-day period using Global Positioning System (GPS) tracking in Sweden. However, the longest dispersal distance in a straight line was 294km which occurred in a 22-day period (Walton et al., 2018). The size of red foxes' territories and the

reasons they disperse are varied. The influencing factors of dispersal in red foxes seem to depend on population densities and size of home ranges occupied. Male red foxes and particularly sub adult males were found to be more likely to disperse large distances. Nevertheless, long distance dispersal is rare in red foxes, but it does show that genetic diversity may occasionally be exchanged between populations (Walton et al., 2018).

Goszczyński, (2002) argues the smaller the home range of an individual or group of red foxes the more territorial they are, and so large home ranges overlap with less aggression. The size of the male (Iossa, Soulsbury, Baker, & Harris, 2008), altitude and proportion of agricultural land (Walton, Samelius, Odden, & Willebrand, 2017) seems to influence territory size. However, Goszczyński, (2002) and McDonald, (1979) argues the main reasons determining the size of a red fox territory is related to food distribution and availability. Furthermore, Holmala & Kauhala (2006) suggests that home ranges on the larger side are usually located in barren habitats such as heaths while the smallest home ranges are in urban areas. Home range size varies from a few dozen hectares (Goszczyński, 2002) to 358 square kilometres (Walton et al., 2017).

Red foxes generally use a smaller home range when living in urban areas and so reach much higher densities frequently exceeding rural population densities. This difference is at least partially due to the rich bounty of food that urban living provides (Goszczyński, 2002; Luniak, 2004; Mackenstedt, Jenkins, & Romig, 2015). Goszczyński (2002) argues the larger the home range the less conspecific interactions take place and so the spread of diseases such as rabies may be less likely to be facilitated.

All in all, urban red foxes tend to be more sedentary in their habits than their rural living conspecifics (Luniak, 2004). Additionally, Wandeler, Funk, Largiadèr, Gloor, & Breitenmoser (2003) found in their study that the genetic differentiation between two rural

populations was 0.9% while the genetic differentiation between rural and urban population was between 2.4-5.4%. The relatively large difference in genetic differentiation suggests that urban and rural populations of red foxes may be distinct populations with little gene flow between them. Furthermore, the genetic diversity was higher in the rural red fox populations than in the Zurich city population (Wandeler et al., 2003). A disparity between males and females in how they disperse has been mentioned but this difference between the sexes in red foxes also extends to other aspects of life.

1.7 Red Fox Ecology

Red foxes exhibit sexual dimorphism and vary in length and weight across their range with males generally weighing more (4.3 to 7.6kg) and measuring longer (96 to 115cm) than females (3.6 to 6.5kg and 91 to 110cm long respectively) (Cavallini, 1995; Pagh, Hansen, Jensen, Pertoldi, & Chriél, 2018). Although the validity of Bergman's rule is in question (Yom-Tov, Benjamini, & Kark, 2002), and some species such as some subterranean rodents follow the converse to Bergmann's rule (Medina, Martí, & Bidau, 2007), red fox populations living higher in latitude (and colder areas) are usually larger following Bergman's rule just as around 65% of other mammals seem to do so (Sablin & Germonpre, 2004).

Variation and flexibility in a changing environment has likely contributed to the red fox's success to become the most widespread carnivore in the world (Gomes and Valente, 2016; Statham *et al.*, 2014). Red foxes are particularly adaptable, living in a wide range of habitats. These habitats include, boreal forests, temperate forests, alpine environments, arctic tundra, hot dry deserts, meadows, and urban areas (Cagnacci, Lovari and Meriggi, 2003; Careau, Morand-Ferron and Thomas, 2007; Jędrzejewski & Jędrzejewska, 1992; Kurki, Nikula, Helle, & Harto, 1998; Vuorisalo et al., 2014).

Within these habitats red foxes in Europe live in a variety of altitudes from sea level to at least 1350m above sea level (Malczewski, Gawor, & Malczewska, 2008) and the species is known to exist >2300 m above sea level in North America (Swanson, Fuhrmann, & Crabtree, 2005). Different habitats produce different challenges with variations in prey, competitors, predators, and parasites in addition to the climatic and physiological variables encountered. For example, in North America populations of montane red foxes apparently differ from red fox's native to other close by regions in ecological and morphological characteristics, and even differ phylogenetically (Sacks, Statham, Perrine, Wisely, & Aubry, 2010).

The speed to which local adaptations occur depends on the genetic variation and strength of the selection coefficient. Man can exert a strong selective pressure, for example, in the case of desired traits bred in foxes in just 10s of generations (Kukekova et al., 2018). Following the assumption of a 1-year generation time as Sacks *et al.*, (2010) it seems with very high selective pressures certain traits or local adaptations may be produced in only 40-50 years (Kukekova et al., 2018) in rural or urban areas.

Urban areas are the most recent of the habitats that have been incorporated into the red foxes list of terrains. Covering an area of around three percent of the earths land surface this represents a new habitat to wildlife with its own challenger's (Johnson & Munshi-South, 2017). Expanding cities have been met with a decline in biodiversity yet there are some species such as the red fox that exploit this to their advantage. Cites contain concentrated food sources which brings more foxes together which as a consequence may also help facilitate the spread of disease (Goszczyński, 2002; Steck and Wandeler, 1980).

Diseases can also have a large selective pressure on wildlife (McCallum, 2012) and the red fox is host to many pathogens (of which some are potentially fatal) that may be transmitted to

and from humans and animals including pets (especially dogs) (Letková *et al.*, 2006; Plumer *et al.*, 2014). For example, a study by De Liberato *et al.*, (2018) examined 262 individual faecal samples over a 12 month period from stray dogs taken into a shelter in Central Italy. Here they found twelve different parasitic taxa infections of which many also affect foxes including *Giardia duodenalis*, *Toxocara canis*, *Trichuris vulpis*, *Angiostrongylus vasorum*, and *Crenosoma vulpis* (Deak *et al.*, 2017; Debenham *et al.*, 2017; Jeffery *et al.*, 2004; Karamon *et al.*, 2018). In Great Britain another study was done over a 13 month period by Smith *et al.*, (2003) which used 586 faecal samples from red foxes. This study also found a multitude of parasites including *Taenia pisiformis*, *T. canis*, *T. vulpis*, *T. leonine*, *Dipylidium caninum*, and *Uncinaria stenocephala*.

T. canis is an intestinal ascarid and is commonly found in both dogs and red foxes (definitive hosts) in Europe which it can transmitted to humans and each other through environmental faecal contamination or the ingestion of infected paratenic hosts such as rodents, birds and invertebrates (Reperant et al., 2007). Arvicolid rodents are a favoured prey item for red foxes, these rodents act as the main intermediate hosts in Europe for the potentially human fatal (in humans and other animals) *Echinococcus multilocularis* tapeworm while the red fox acts as the definitive host (Laurimaa et al., 2015).

Some pathogens like *Toxoplasma gondii* seem to have coevolved with red foxes and other animals over a long time and cause little known harm. Other pathogens infecting red foxes on the other hand, such as the rabies virus, seem to be more recent and may heavily reduce their populations (Rupprecht, Turmelle, & Kuzmin, 2011b).

The negative-strand RNA virus that causes rabies in foxes, other animals, and humans belongs to the Family Rhabdoviridae and Genus *Lyssavirus*. The primary niche of these RNA viruses is the central nervous system of their mammalian hosts especially Chiropteran bats

and several species of Carnivora (Rupprecht, Turmelle, & Kuzmin, 2011a). Usually a lyssaviruses and their specific host reservoir population can adapt to one another but if this lyssavirus infects a different population a fatal self-limiting rabies-like infection typically occurs (Bourhy et al., 1999a).This seemed to have happened when the rabies virus started infecting red foxes which trailed a decrease in incidence among urban dogs and wolves (Bourhy et al., 1999b). This may have occurred as foxes occasionally eat bats and bat lyssaviruses to carnivores does occur (Rupprecht et al., 2011b). This infection had high death rates and seems to have kept red fox populations abundance in check with arears of higher fox abundance being reduced most as it is also transmitted by fox bites to conspecifics (Steck & Wandeler, 1980b). By the early 1940s rabies-infected foxes started to appear at the former Russian–Polish border before spreading to the rest of Europe (Bourhy et al., 1999b). This suggests adaptations have occurred either with the red fox's immune system or the virus is adapting to be less virulent. However, the adaptation process seems to still be occurring as the introduction of a rabies vaccine in red foxes significantly increased their populations in continental Europe (Goszczyński, Misiorowska, & Juszko, 2008). While infectious mites (*Sarcoptes scabiei*) that causes scabies also had a large impact, for example by depleting the Nordic red fox populations by over 70% (Forchhammer & Asferg, 2000).

Parasite-host relationships are a natural part of the ecosystem that may regulate population levels of their hosts (Albon et al., 2002) and plays a critical role in evolution (Møller & Szép, 2011). Living in urban areas may cause a shift in disease dynamics and with abundant food red foxes can make an easier livelihood (Bradley and Altizer, 2007; Reperant *et al.*, 2007; Vuorisalo *et al.*, 2014).

The red fox is a generalist predator taking all sorts of food both animal and plant based. Red foxes search for prey by sight, smell and by listening for tell-tale signs. They are opportunistic in their feeding habits and their diet depends on many variables. Their diet is

usually predominately composed of small mammals but as seasons change so does the abundance and availability of food items such as rodents. Red foxes respond to this seasonal change sometimes by changing foraging behaviour (Kidawa & Kowalczyk, 2011). For example, in certain months of the year foxes may consume earth worms amounting to over 60% of their calorie intake (Macdonald, 1980).

The food red foxes consume is dependent primarily upon habitat type (Kidawa & Kowalczyk, 2011) and prey availability which includes small rodents, insects, birds, lagomorphs, fruits, carrion, eggs, frogs, reptiles, and even fish, (Dell'Arte, Laaksonen, Norrdahl, & Korpimäki, 2007; Kauhala, Laukkanen, & Rége, 2007; Macdonald, 2010).

The diet of urban red foxes and that of rural red foxes often differs. For example, in central Europe the greater part of a rural fox's diet is composed of rodents from the Arvicolidae family while in urban foxes' anthropic foods, more invertebrates and less rodents are more common (Reperant et al., 2007). Dispersal between these areas is possible, although in some areas urban and rural fox populations may remain genetically isolated from one another (Wandeler et al., 2003).

1.8 Objectives:

Although the red fox has been studied since the 1930s (Harris & Rayner, 1986) there seems to be no genome-wide study involving wild populations across the European continent (and one population in Siberia, Asia) which searches for signals of selection for this species at the time of writing. The access of tools proven to successfully work in genome-wide analysis of SNPs in domestic dogs has allowed much potential for further exploration in population genetic studies for other wild canids (Karlsson *et al.*, 2007; Lindblad-Toh *et al.*, 2005). This approach has been used in several wild canid studies (Pilot *et al.*, 2014; Schweizer, *et al.*,

2016; VonHoldt *et al.*, 2016; Zhang *et al.*, 2014) which also includes semi-domesticated red fox studies (Johnson *et al.*, 2015; Kukekova *et al.*, 2018) to understand variants in the genome associating with some adaptive trait. These studies have uncovered traits in the genome relating to size, coat colour, metabolism, behaviour, vison, hearing and more.

To address this gap in knowledge of the wild red fox I used GBS data (provided by A. McDevitt, *unpublished data*) from over 500 individuals across 22 countries in Europe (and one location in Siberia). The purpose of this study was to:

1) ascertain population structure
2) to identify local adaptations within Europe (Eurasia), and
3) uncover potential genomic signals of selection that may occur such as genes related to immunity and metabolism.

The approaches I took to identify SNPs showing signals of selection were composed of the following steps. First, I identified SNPs that had an outlier allele based on the genomic data statistics by using the q-value in the PCAdapt package (Luu et al., 2017). The range expansion model was used based on the recent phylogeographical evidence of this species. Next, BayeScan (a model-based method) was used to identify SNPs that are significantly differentiated among populations, strengthening the evidence for diversifying selection (Foll & Gaggiotti, 2008). Third I calculated allele frequencies for the candidate loci/genes exhibiting diversifying selection by implementing an allele frequency analysis in adegenet R-package (Jombart, 2008). The candidate genes were also analysed for significant enrichment of Gene Ontology (GO) category terms (Huang, Sherman and Lempicki, 2009; Reimand *et al.*, 2016). These GO terms give information on broad functional characteristics of the candidate genes found. Subset or ‘child’ GO terms give more specific functional characteristics of the genes. The GO ‘child’ terms were then looked at (Binns *et al.*, 2009;

Shimoyama *et al.*, 2015) before using Google scholar to find detailed studies on these candidate genes in the scientific literature. This information was used to generate hypotheses on possible selective pressures that may have been the cause of local adaptations in the red fox.

2. Methods

2.1. Study Area

I started with GBS information taken from 561 red fox samples (one from each individual fox) (provided by A. McDevitt, *unpublished data*). These samples (Figure 1a) came almost exclusively from across the European continent (Figure 1a) barring the countries that were exhibiting concerning issues with the rabies virus. However, 54 samples were also taken from individuals in eastern Siberia which is situated in the Asian continent (Figure 1a).

Figure 1a Fox samples shown across Europe (and Siberia) before filtering. Figure produced by using (QGIS Developmental Team, 2016).

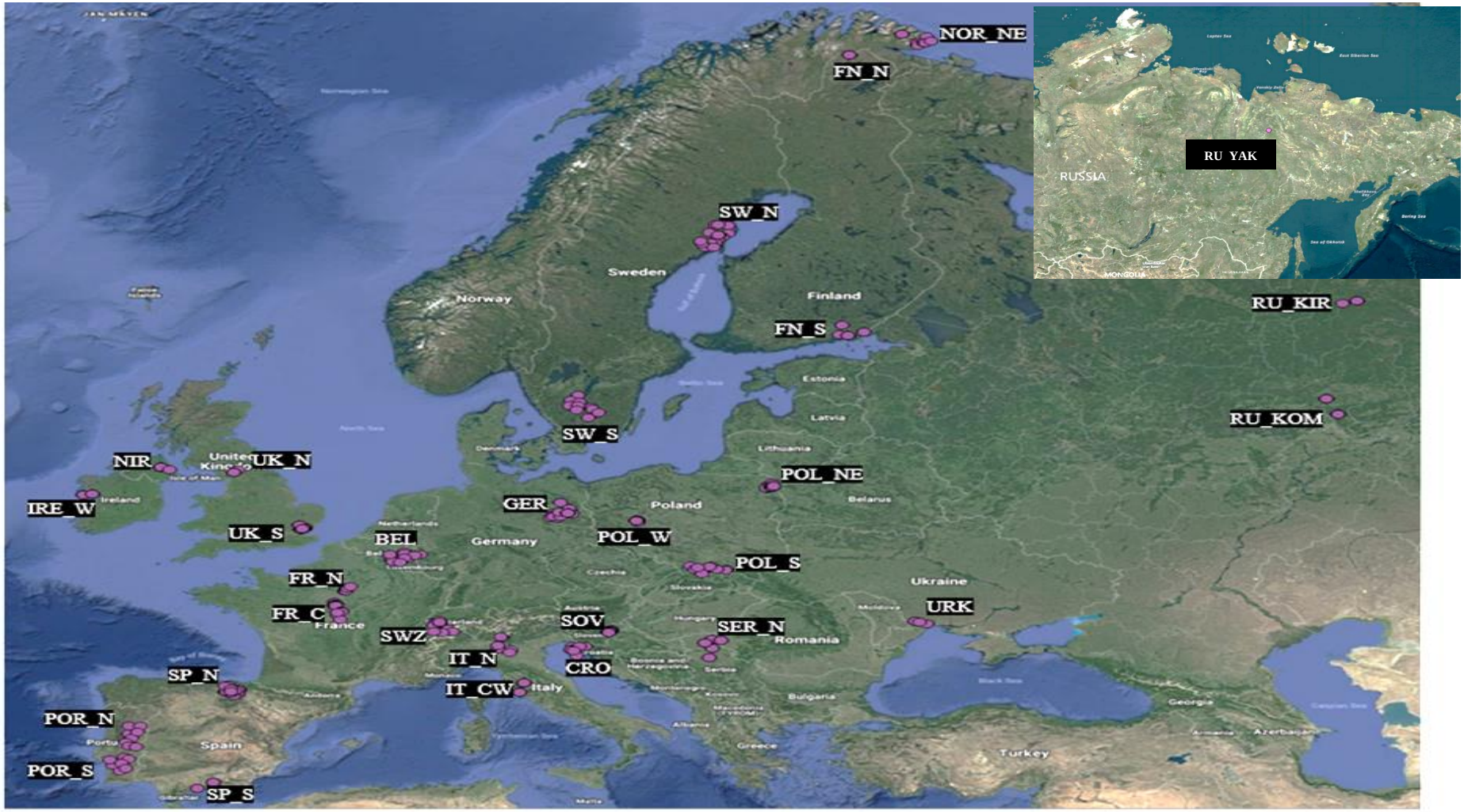

Figure 1b Fox populations after filtering. The 29 red fox populations sampled across Europe are shown distributed across 19 countries and the Siberian population RU_YAK is shown in the inset. Each country contains one to three populations and each population contains seven to 28 individuals which are shown in Table1. Figure produced by using (QGIS Developmental Team, 2016).

Table 2 The populations are designated by code which identifies the location of the population which is followed in brackets by the number of individuals within that population.

Continent	Country	Number of Individuals	Number of Populations	Populations
Europe	Belgium	**14**	1	BEL(14)
Europe	Croatia	**9**	1	CRO(9)
Europe	Finland	(8+12)	2	FN_N(8)
		20		FN_S(12)
Europe	France	(7+13)	2	FR_N(7)
		20		FR_C(13)
Europe	Germany	**28**	1	GER(28)
Europe	Ireland	**20**	1	IRE_W(20)
Europe	Italy	(7+20)	2	IT_N(7)
		28		IT_CW(20)
Europe	Northern Ireland	**20**	1	NIR(20)
Europe	Norway	**15**	1	NOR_NE(15)
Europe	Poland	(16+13+19)	3	POL_NE(16)
		48		POL_S(13)
				POL_W(19)

Europe	Portugal	(7+8)	2	POR_N(7)
		15		POR_S(8)
Europe	Russia	**27**	2	RU_KOM(8)
				RU_KIR(19)
Asia	Russia	**9**	1	RU_YAK(9)
Europe	Serbia	**18**	1	SER_N(18)
Europe	Slovenia	**25**	1	SOV(25)
Europe	Spain	(26+7)	2	SP_N(26)
		33		SP_S(7)
Europe	Sweden	(16+10)	2	SW_N(16)
		26		SW_S(10)
Europe	Switzerland	**19**	1	SWZ(19)
Europe	England	(9+27)	2	UK_N(9)
		36		UK_S(27)
Europe	Ukraine	**7**	1	URK(7)

NOR_NE = Norway northeast, SW_N = Sweden north, SW_S = Sweden south, FN_N = Finland north, FN_S = Finland south, UK_N = England north, UK_S = England south, IRE_W = Ireland west, NIR = Northern Ireland, FR_N = France north, FR_C = France central, SP_N = Spain north, SP_S = Spain south, POR_N = Portugal north, POR_S = Portugal south, BEL = Belgium, GER = Germany, SWZ = Switzerland, IT_N = Italy north, IT_CW = Italy central western, CRO = Croatia, SOV = Slovenia, POL_NE = Poland northeast, POL_S = Poland south, POL_W = Poland west, SER_N = Serbia north, URK = Ukraine, RU_KOM = Russia northeast, RU_KIR = Russia central east, RU_YAK = Russia northeast. These populations are the same as in Figure 1b.

2.2. Laboratory Methods

Red fox tissue samples were obtained from freshly culled, frozen, ethanol (70-95%) and DMSO-preserved material. Genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) according to manufacturer's protocols (with the additional treatment of RNase). DNA was quantified on a Qubit™ 3.0 Fluorometer (ThermoFisher Scientific) with a Qubit™ dsDNA BR (Broad Range) Assay Kit. A total of 100ng/µg of high molecular weight DNA (visualized on an electrophoresis gel) was sent to the Genomic Diversity Facility in Cornell University (USA) where Genotype-by-Sequencing (GBS) was used for constructing reduced representation libraries (Elshire et al., 2011). Briefly, the restriction enzyme EcoT22I (ATGCAT) was used to digest the DNA of individuals in six GBS libraries (each consisting of 95 uniquely barcoded individuals and one negative control). Individual ligations were then pooled and purified from adaptor excess using the QIAquick PCR Purification Kit (Qiagen). For library preparation genomic adaptor-ligated fragments were then PCR amplified in 50 µL volumes with 10 µL of the DNA fragment pool, 1 × Taq Master Mix (New England Biolabs Inc.), and 12.5 pmol each of the following Illumina primers:

5´AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT and

5´-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT (the underlined parts will hybridize to the two Illumina flowcell oligos). Temperature cycling consisted of 72 ºC for 5 min, 98 ºC for 30 s followed by 18 cycles of 98 ºC for 30 s, 65 ºC for 10 s, and 72 ºC for 30 s, with a final extension step at 72 ºC for 5 min. The EcoT22I GBS libraries (now containing ID tags and Illumina flowcell adaptors) were purified again as described above. An aliquot was run on the BioAnalyzer™ 2100 to verify

fragment sizes and for the removal of adapter dimers. Library DNA was then quantified on a Nanodrop 2000 (ThermoFisher Scientific) and subsequently sequenced on an Illumina HiSeq 2000 (Cornell University, Life Sciences Core Facility).

2.3. Population Criterion, Data filtering, and SNP calling

The raw Illumina sequence data from 561 individuals (two individuals were repeated after initial poor quality) were processed through the GBS analysis pipeline implemented in TASSEL v5.2.31 (Trait Analysis by aSSociation, Evolution and Linkage) (Glaubitz et al., 2014). TASSEL is a software Package that evaluates evolutionary patterns, traits associations, and linkage disequilibrium. TASSEL can also handle indels (insertion & deletions) with ease. A total of unique 9,249,177 tags were found. Because of concerns about the performances of *de novo* approaches to identify SNPs in reduced representation genomic techniques, the reference genome of the most closely related canid species available, the domestic dog (CanFam3.1; *Canis lupus familiaris*) was used to align the sequence tags and determine the location of said tags on individual chromosomes using the Burrows-Wheeler alignment tool (H. Li & Durbin, 2009). A total of 6,267,895 tags (67.8%) were uniquely aligned to the dog genome, and the rest of the tags were disregarded from further analyses. Identified SNPs were initially filtered with a minor allele frequency (MAF) of <0.01 and missing data per site <0.9, resulting in 144,745 SNPs with a mean individual depth of 15.853 (±SD = 4.629). This initial work was done by Cornell University in the Buckler Lab (Cornell University, Life Sciences Core Facility).

This original initially filtered data (provided by A. McDevitt, *unpublished data*) was within a Variant Call Format (VCF) file which contained 561 samples and 144745 variants. This data set also had the accompanying Sequence Alignment/Map (SAM) file (see below). An excel

sheet with the longitudinal and latitudinal coordinates of the samples was uploaded to the Quantum Geographic Information System (QGIS) software (QGIS Developmental Team, 2016). QGIS allows spatial information to, be displayed on a computer, create maps, and can also perform spatial analysis. QGIS was used in this study to enable the samples to be sorted according to the population criterion used (see below) and to be seen on a visual world map (Figure 1b).

By using QGIS and the sample's coordinates, the distance between the samples was calculated by inputting the coordinates into a longitude and latitude calculator (Veness, 2018). Based on the red foxes' dispersal rates and ecology the samples that were collected outside an 80km radius of any cluster was removed using TASSEL 5.2.43 (Bradbury et al., 2007). This is the basis for the population criterion. After removing any samples that did not conform to the population criteria a VCF file containing 469 samples with 144745 variants was produced.

This new file was loaded back into TASSEL 5.2.43 (Bradbury et al., 2007) and after each command was subsequently saved as a VCF file which was loaded for the next step of filtering. TASSEL was used to remove sites that had indels and to change the third state to missing before imputing by Beagle v.4.

This resulted in a file containing 469 samples and 107192 variants. Individual and SNP quality filtering was then performed again using TASSEL 5.2.43 (Bradbury et al., 2007). The number of loci was reduced in TASSEL via the option 'sites minimum count'. SNPs with an observational reading of 20% or more missing data (equivalent to - -geno 0.2 in Plink) was removed, leaving 469 samples and 40443 variants.

The next step was to remove any individual with more than 20% missing data in which the sample number was reduced via the option 'minimum proportion' (equivalent to - -mind 0.2

in Plink), leaving 443 samples and 40443 variants. Further SNP filtering in TASSEL via the option ‘minimum frequency’ (equivalent to --maf 0.01 in Plink, see below) reduced the number of loci of individual SNPs with a minor allele frequency (MAF) of less than 1%, leaving 436 samples and 38907 variants.

This new data set was then transferred to Plink v1.90b5 (Purcell et al., 2007) and the names of the populations were assigned. The renaming was done in Plink v1.90b5 by using the commands in the section ‘Update Individual Information’ within the PLINK 1.9 manual (Purcell et al., 2007). Plink v1.90b5 was then used to perform a LD test with the option ‘indep-pairwise 50 5 0.7’ (50 SNP window size, shifting the window 5 SNPS at each step, 0.7 variance inflation factor (VIF)). This in-depth pairwise LD process works by looking at a block of 50 paired SNPs and compares LD between each pair of SNPs. A pair of SNPs will be removed if the LD is greater than 0.7 and then the window moves down 5 SNPs and the process will be repeated. This form of quality control filtering reduces genotyping error. This kept the number of samples at 436 and pruned the loci to 36508 produced in a new file.

This new data set was then converted into a genepop file (gen) via PGDSpider 2.1.1.5 (Lischer & Excoffier, 2012). In the data output file option file format: GENEPOP was chosen. A SPID file was created in the SPID conversion script.

This genepop file has the data arranged into populations based on the population criteria mentioned above. This file was uploaded into Rstudio version 1.1.423 (RStudio Team., 2016). A Hardy-Weinberg test (HWE) test was then performed per population using the package Pegas (Paradis, 2018) in conjunction with the adegenet and ape packages (Jombart, 2008; Paradis & Schliep, 2019). Pegas (Population and Evolutionary Genetics Analysis System) amongst other functions plots, analyses, reads, writes, and manipulates allelic and haplotypic data. Pegas also analyses population nucleotide sequences for population structure

(FST) and equilibrium (HWE) (Paradis et al., 2018). An HWE test was ran with a bootstrap value of 1000. The HWE tests whether, through a series of loci, genotype frequencies follow or deviate from the Hardy–Weinberg Equilibrium hypothesis (Paradis et al., 2018).

The HWE alpha value was originally set at 0.05 (representing the p-value) and then corrected by using the bonferroni correction test (0.0001) (da Silva Santos *et al.*, 2012; Khatkar *et al.*, 2008; Purfield *et al.*, 2012) to reduce the false positive rate (FDR). If the p-values for a particular SNP in the HWE test were not statistically significant (>0.0001) in over half of the populations of that particular SNP, the SNP was removed. There were 5800 SNPs that were removed due to the HWE test. These SNPs were added to note pad and saved as a .txt file. The SNPs that were to be removed was removed by Plink v1.90b5 (Purcell et al., 2007) via following the section 'Extract a subset of SNPs: file-list options' in the manual. This produced a file containing 436 samples and 30708 variants. This was the filtered data set upon which the following analysis took place.

2.4. Detecting Outliers

To identify the genomic outlier loci, I applied two methods with each one having a different underlying statistical methodology. First, BayeScan was used, as many other population genetic studies have used (Fischer *et al.*, 2014; Freedman *et al.*, 2010; Schweizer, *et al.*, 2016), to look for signals of selection and in which to identify what type of selection has taken place. The 30 populations above (Table 2) was used within the implementation of the BayeScan analysis while PCAdapt has no need for prior information regarding populations. PCAdapt has also been used in many population genetic studies (Arnyasi *et al.*, 2017; Basto *et al.*, 2016; Ji *et al.*, 2012) to ascertain population groupings.

BayeScan is a Bayesian method and works on the multinomial Dirichlet model (Foll & Gaggiotti, 2008). BayeScan looks at differences in allele frequencies between different sub-populations or demes within a single gene pool. It then tests whether any of these differences in allele frequencies are specific to one deme and if they significantly differ from other demes within the collective gene pool. The allele frequencies are measured by an F_{ST} coefficient and a posterior probability (alpha) is allocated to a model. In this assigned model a difference in allele frequencies is better explained by natural selection than by the null model (Schweizer, et al., 2016). The alpha value can be positive or negative which implies diversifying selection or balancing selection is respectively taking place. Negative alpha values suggest balancing selection (Foll & Gaggiotti, 2008). As a larger number of populations gives better results the balancing selection was ignored in this study. Positive alpha values suggest diversifying selection which overall suffers less false positives and requires a smaller number of populations to obtain more reliable results compared to the balancing selection criteria (Lotterhos & Whitlock, 2014). It is diversifying selection that will be focussed on in this study. In BayeScan prior odds was run at the default 10 in order not to discard true loci under selection (Lotterhos & Whitlock, 2014). To reduce the false positive rates a false discovery rate (FDR) of 0.05 was used even though this increases the risk of signals of selection being missed (Foll & Gaggiotti 2008). To reduce the false positive rate, of identifying outliers of natural selection, even further, PCAdapt was independently run on the same data set (see below).

PCAdapt was used to detect unexpectedly large changes in allele frequencies between populations. This indicates an environmental change has occurred in which natural selection may have responded (Luu et al., 2017). First, Principal Component Analysis (PCA) was implemented through R-studio via the package PCAdapt (Luu et al., 2017) and visually

modified (colours and shapes) by ggforty and ggplot2 (Tang, Horikoshi and Li, 2016; Wickham., 2016) using the same filtered data set (ped file) as was used in BayeScan. This positioned samples into a hierarchical population structure. The null assumption from PCAdapt is that no SNP is under selection. Population differentiation in PCAdapt is estimated at each SNP by comparing the SNPs value with an expected estimate, determined from the genome background as a whole (Dalongeville et al., 2018). By plotting scree plots the optimal number K of principle components (PCs) to use can be determined. This will give the difference in variance in a percentage which is explained by each PC. I used the range expansion model to take into account the red foxes life history and general ecology which includes living during glacial and interglacial periods (Luu et al., 2017; Sommer and Benecke, 2005; Statham *et al.*, 2018). I performed a scree plot, although there are some controversies it seems that a scree plot can be presumed to produce reasonably accurate results (Luu et al., 2017; Zwick & Velicer, 1982). I followed Luu et al., (2017) by using Cattell's rule and kept the PCs that correspond to eigenvalues to the left of the straight line on the scree plot, which in this case was three. Cattelll's Rule states that random variation is represented by the eigenvalues on the strait line while the eigenvalues that depart from this line represent population structure (Luu et al., 2017).

PCAdapt was also used to detect outlaying SNPs using a statistical test called the Mahalanobis distance. The Mahalanobis distance measures the distance between K correlations of the SNP and each axis and mean correlations. When this is scaled by a constant a chi-squared distribution with K degrees of freedom a calculation will be produced with the null hypothesis that there are no outliers (Luu et al., 2017). The default with SNPs of P-values with a MAF smaller than 0.01 are not computed on a log10 of the p-values when a Manhattan plot is produced (Luu et al., 2017). I modified the Manhattan plot with the ggman package in R-studio to display the results in a clearer way (Figure 5). Using the FDR (q-

value) is recommended by Luu et al. (2017) to produce outliers. An FDR was used in this analysis at 0.05 to have a compromise between detecting false positives and missing loci under selection. The cut-off point for detecting outliers was chosen to be 0.05 to be consistent with BayeScan's analysis (see above). An FDR value of 0.05 has been used in previous studies (Buonaccorsi *et al.*, 2017; Josephs *et al.*, 2019).

Both BayeScan and PCAdapt have strengths and weaknesses, whereas BayeScan requires prior information on the samples populations PCAdapt arranges the samples at the individual level without possible population bias, hence these two methods complement one another at different levels (Coop *et al.*, 2010; Foll and Gaggiotti, 2008). The FDR when using admixed individuals is relatively high in BayeScan but is compensated by PCAdapt having a relatively low FDR (Luu *et al.*, 2017). By crossing the outliers, we get information from the population level and the individual level while also minimizing the FDR.

2.5. Identifying Candidate genes

The SNPs from the analysis suggesting diversifying selection has occurred was then blasted to find the location on the genome. The search began by finding the up to 64 nucleic bases in the accompanying Sequence Alignment/Map (SAM) file with the unique SNP code found in the initially filtered VCF file (see above). A SAM file is a TAB-delimited text format that contains essential information such as mapping positions (Li et al., 2009). This information allows the location of specific alleles to be found on the aligned genome. The stretch of nucleic bases or Deoxynucleic acid (DNA) that surrounded the SNPs that had been identified as candidate loci for selective processes from the methods above was copied and pasted then blasted into the Basic Local Alignment Search Tool (BLAST) (Kerfeld & Scott, 2011). As the information used to put into BLAST was nucleotides, I used BLASTn. In the BLASTn

settings the options were changed in database to RefSeq Genome Database (refseq_genomes), in organism to *C. lupus familiaris* (taxid:9615) and the other options was left as default ("BLAST: Basic Local Alignment Search Tool", 2019). Then in the box Enter Query Sequence I pasted the nucleotide sequence that was taken from the SAM file. This gave the location in the genome and if it landed within a gene (and exon or intron) or not. The criterion used here was if the SNP was found outside of a gene then it would not be used. Furthermore, short nucleotide sequences that produced multiple hits were also disregarded. These criteria were used to strengthen the likelihood of obtaining more reliable results.

The significance of the input nucleotide sequence which matches the database sequence is measured by the Expect (E) Value. The lower the E values the more significant the match and the more reliable the result (Kerfeld & Scott, 2011).

2.6. Gene ontology and Enrichment

The candidate genes were further analysed for significant enrichment of GO categories using g:Profiler (Reimand *et al.*, 2007; Reimand *et al.*, 2016). I entered my candidate gene list and selected *C. lupus familiaris*. Next I used the same parameters as Schweizer *et al.*, (2016), I scrutinised significant categories ($P \leq 0.05$) with a minimum of two genes after the correction for multiple testing using the Benjamini–Hochberg FDR was completed.

I also independently used another functional analysis tool, the Database for Annotation, Visualization and Integrated Discovery (DAVID). Like g:Profiler, DAVID maps a set of genes to associated biological annotation terms such as the GO terms and KEGG pathways used here. Next the gene members are enriched by statistically inspecting theses terms against a background resulting in the identification of associated (enriched) terms with gene members (Huang, Sherman and Lempicki, 2009). While g:Profiler is a set of class-1 category enrichment tool David is both calss-1 and class-3, the differences are summarized in Table 3 (Huang, Sherman and Lempicki, 2009). In the annotation summary results the background was selected for *C. lupus familiaris*, all 24 of the Gene_Ontology and only the KEGG_Pathway boxes were selected. In the Functional Annotation Clustering combined view section the default parameters were kept with an exception of the Kappa similarity threshold which was changed to 0.65 to give a substantial agreement of enrichment (Viera & Garrett, 2005). I choose to focus on the GO terms with an enrichment of >1.3 which is equivalent of a <0.05 P-value (ease score). For a summary on these parameters see Table 4.

Table 3. Modified from Huang, Sherman and Lempicki, (2009) showing how class-1 and class-3 enrichment tools work.

Tool category	Description	Limitations and indication	Sub-type of algorithms	Methods	Example tool
Class 1: Singular Enrichment Analysis (SEA)	The input gene list is used to calculate an enrichment P-value for each term. This is followed by the enrichened terms being listed in a simple liner text format. This is the usual traditional algorithm and is still the predominant enrichment analysis tool used.	Can analyse any gene list, which can be selected from any high-throughput biological studies or technologies. For example Micoarry,ChIP-on-sequence, SNP array, EXON array etc. However, the depth of the interrelationships among the terms may not be fully captured in linear format report.	Global reference background	Fisher's exact hypergeometric chi-square binomial	GoStat, GOtoolBox, GFinder, GARBAN, FatiGO, etc. BayGO **DAVID**
			Local reference background	Fisher's exact hypergeometric	**g:Profiler**
Class 3: Modular Enrichment Analysis	This strategy incorporates the basics of SEA. However, the calculation of the P-values also includes the term-term/gene-gene relationships. This provides the advantage of the term-term/gene-gene relationship may contain unique biological meaning that is not held by a single gene or term. This modular/network analysis is closer to the natural biological data structure.	As in Class 1, can analyse any gene list, which can be selected from any high-throughput biological studies or technologies. In addition, emphasis is on network relationships during analysis. 'Orphan' gene/term (with little relationships to other genes/terms), that occasionally may be very important, too, may be left out from the analysis.	Composite annotations	Measures the enrichment on joint terms	geneCodis, ProfCom etc.
			DAG Structure	Measures the enrichment by also considering parents-child relationship	topGo, POSOC, etc.
			Global Annotation relationship	Measures term-term global similarity with Kappa statistics Czekanowski-Dice Person's correlation	**DAVID,** GoToolBox etc.

Table 4 Modified from Huang, Sherman and Lempicki, (2009) showing a summary on how the statistics are used in in enrichment analysis methods.

Module /page	Statistics/ parameters	Explanation/definition	How to understand the value
Gene IDs Used	Background genes or population genes	To choose the percentage of enrichment, a certain background must be set up to be compared with the user's gene list. For example, 10% of user's genes are kinases whereas only 1% of genes in human genome (this is population background) are kinases. Thus, in this example the conclusion is clear that the user's study is highly related to kinase. However, this 10% itself cannot provide such a conclusion without comparing it with the background information	Generally, the reference background is used in the context of the pool of genes that have a chance to be selected for the studied annotation category under the scope of users' particular study Default background is the entire genome-wide genes of the species matching the user's input IDs. Previously built backgrounds, for example, genes in Affymetrix chips etc, are available for the user's choice in principle, smaller P-values usually occurs with a larger gene background. Almost all of the high throughput studies are, or at least are close to, genome-wide scope, for regular cases in general the default background is usually good enough.
Functional classification of genes	Stringency of classification	To control the behaviour of DAVID Fuzzy clustering	A general guideline is to choose higher stringency settings for tight, clean and smaller numbers of clusters and lower for loose, broader and larger numbers of clusters. A level can be chosen to suit the requirements needed which starts at the Default medium.
	Each group's Enrichment score	Ranks overall importance(enrichment)of gene groups. By using the geometric mean an enrichment P-values (EASE scores) for each annotation term associated with the gene members in the group is produced. minus log transformation is applied on the average P-values to emphasize that the geometric mean is a relative score instead of an absolute P-value.	The higher the score the more important (enriched) the group of terms of the gene members are in a given study; therefore, more attention should go to them. Enrichment score of 1.3 is equivalent to non-log scale 0.05. Therefore, more attention should be given to groups with scores Z1.3 However, the gene groups with lower scores may be explored which may provide additional information.
Functional Annotation Chart1	P-value/ EASE score	To examine the significance of gene–term enrichment with a modified Fisher's exact test (EASE score). For example, 10% of user's genes are kinases versus 1% of genes in human genome (this is population background) are kinases. Thus, the EASE score is o0.05, which suggests that kinases are significantly more enriched than random chance in the study for this particular example	The smaller the P-values, the more significant they are. Default cut off is 0.1 but it can be set to different levels of cut off via the option panel on the top of result page. Due to the complexity of biological data mining of this type, P-values are suggested to be treated as score systems, i.e., suggesting roles rather than decision-making roles. The results should be judged based on how much biological sense they make.
	Benjamini	Globally corrects enrichment P-values to control family-wide false discovery rate under certain rate (e.g. r 0.05). It is one of the multiple testing correction techniques such as	The tore terms examined, the more conservative the corrections are. Because of this, all the P-values get larger. If the interesting terms have significant P-values after the corrections, then this is helpful. But as the multiple testing correction

		Bonferroni, Benjamini and FDR. This is provided by DAVI	techniques are known as conservative approaches, it could hurt the sensitivity of discovery if overemphasizing them. Judgment is necessary and could be critical as discussed above in EASE score in Functional Annotation Chart section
Functional Annotation Clustering	Stringency of classification	Controls the behaviour of DAVID Fuzzy clustering	A general guideline is to choose higher stringency settings for tight, clean and smaller numbers of clusters and lower for loose, broader and larger numbers of clusters. A level can be chosen to suit the requirements needed which starts at the Default medium
	Each group's Enrichment score	Ranks overall importance(enrichment)of annotation term groups. It is the geometric mean of all the enrichment P-values (EASE scores) of each annotation term in the group and emphasizes the geometric mean which is a relative score instead of an absolute P-value. A minus log transformation is applied on the average P-values	The higher the score the more important (enriched) the group of terms of the gene members are in a given study; therefore, more attention should go to them. Enrichment score of 1.3 is equivalent to non-log scale 0.05. Therefore, more attention should be given to groups with scores Z1.3 However, the gene groups with lower scores may be explored which may provide additional information.
	Individual term members P-value/ EASE score	Inspects the significance of gene–term enrichment with a modified Fisher's exact test (EASE score). The calculation of the P-value is done is done as above in Functional Annotation Chart section	See above as in the Functional Annotation Chart section
	Benjamini	Globally corrects the enrichment P-values of individual term members. calculations are done as above in the Functional Annotation Chart section	See above as in the Functional Annotation Chart section

2.7. Candidate genes further research

These GO terms are known as ‘parent terms’ which give a broad view of identification while more specific roles can be found by looking at the ‘child terms’ of which there be many (Grossmann, Bauer, Robinson, & Vingron, 2007). This, what I call ‘root of terms’ can be followed down the decedents (child terms) until a specific example of a role or function a specific gene can have, which can be studied in more detail. For example, by taking the term ‘anatomical structure development’ (which is under the broad term ‘Biological Process’) which includes the ‘AFF3’ gene and is found both by g:Profiler and DAVID, the decedent term ‘embryonic hindlimb morphogenesis’ can be found (Figure 2). I call this the ‘Root Approch’. The AFF3 gene (also found in the ‘embryonic hindlimb morphogenesis’ term) has been shown to be expressed during normal mouse limb development and is found to be lacking expression in the dorsal subectodermal region, suggesting Aff3 plays a role in dorsoventral patterning (Feenstra et al., 2012). This process was repeated for each candidate gene found in this study.

The Root approach I have used here was done by using QuickGO (Binns et al., 2009) and the Rat Genome Database (RGD) (Shimoyama et al., 2015) to use the GO enriched terms from g:Profiler and DAVID to find their ‘child terms’. Next, I obtained information for each gene by searching Google Scholar for studies done on particular genes relating to phenotypic effects that may be seen to possibly influence red fox local adaptations. Finally, I used this information to relate to the red fox. As each gene usually affects more than one trait each sample was chosen by what is best known in influencing phenotypes in which previous studies have found strong support relating to natural selection or has reliable studies on the genes and their function.

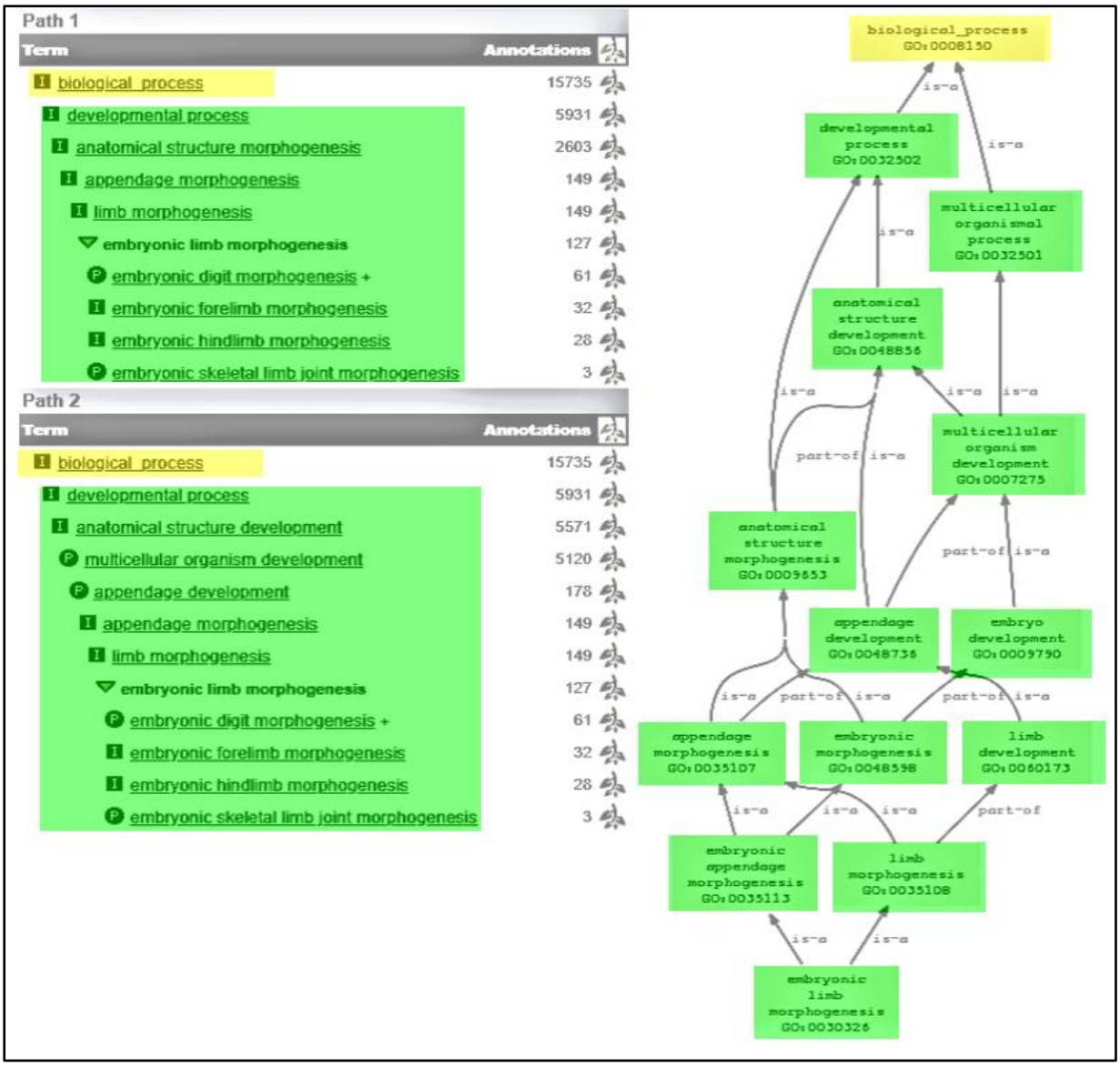

Figure 2 A GO annotation 'root' (right) with the shortest and longest path of ‘child terms’ (green) that can be taken to reach the embryonic hindlimb morphogenesis term (left). (parent terms in Yellow). Figure produced by using http://rgd.mcw.edu.

2.8. Allelic frequency Analysis

The data at this point consisted of SNPs identified as being candidate loci for diversifying selection from the filtering, analyses methods, and reasoning above. To find out the frequencies of these alleles and in what populations, an allele frequency analysis was implemented in R-studio using the package adegenet (Jombart, 2008). Allelic frequencies in multiple populations can uncover signatures of selection and local adaptations in specific populations. This can lead to understanding what the selective pressure/s were and so what caused a particular allele to become more advantageous over others (Beaumont, 2005).

3. Results

After applying the population criterion and bioinformatic filtering the number of samples was reduced to 436 and the number of SNPs to 30,708. This data set is similar in size to the 42,036 data set used in another wide ranging wild canid species (Schweizer, et al., 2016)

3.1. Population Criterion and PCA analysis

By using the population criterion based on ecological data of red fox movements and habits and combining this with information from the location where each sample was taken from, the remaining 436 samples were placed into 30 populations (Figure 1b and Table 2). Of these populations, 29 populations are located in Europe with 427 samples and one population located in east Siberia (Asia) with nine samples (Figure 1b and Table 2). Discrete populations in Europe have previously been described in phylogeographic studies of wide ranging canids (Schweizer, et al., 2016). Furthermore, Statham *et al.*, (2018) argues that red foxes also have distinct populations as do red deer and roe deer which share similar LGM distributions as the red fox (Randi *et al.*, 2004; Skog *et al.*, 2009).

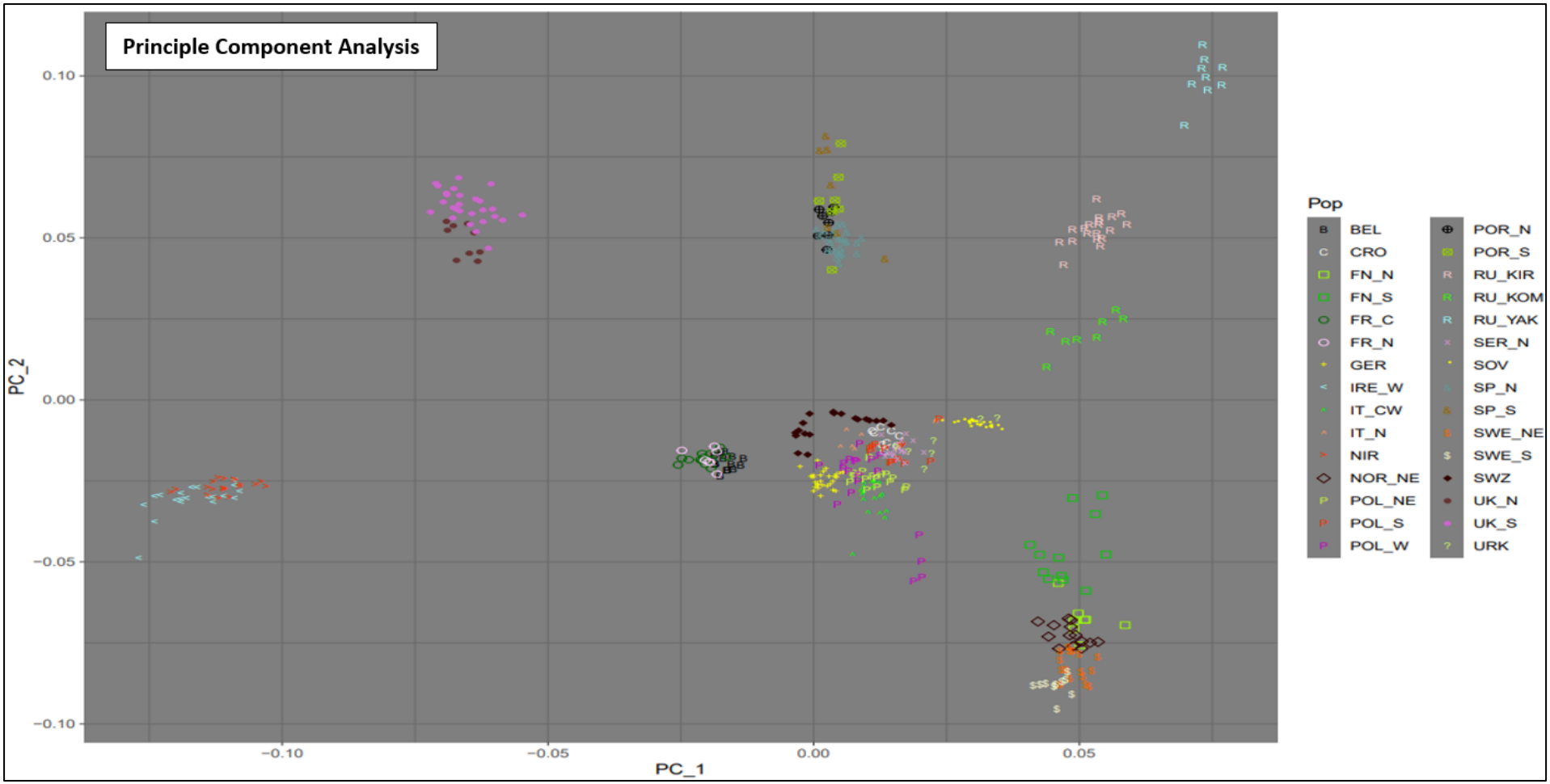

Figure 3 Shows nine groups of the 436 samples used in this study on the first two Principle components. The colour and shape identify each individual relating to the populations assigned using the population criterion (see Methods 2.3). Figure produced by using PCAdapt (Luu et al., 2017) and visually modified (colours and shapes) by ggforty and ggplot2 (Tang, Horikoshi and Li, 2016; Wickham., 2016).

The subsequent PCA implemented using PCAdapt indicated that the populations above clustered into nine groups (Figure 3). The two UK populations (UK_N and UK_S), the two Irish populations (NIR and IRE_W), the four Iberian populations (SP_N, SP_S, POR_N, and POR_S), the five Nordic populations (NOR_NE, FN_N, FN_S, SW_N, and SW_S), the two French and one Belgium populations (BEL, FR_N, FR_C), and the other 11 Central European (and Balkan populations) (CRO, GER, IT_CW, IT_N, POL_NE, POL_S, POL_W, SER_N, SOV, SWZ, and URK), are grouped together while the three Russian populations (RU_KIR, RU_KOM, and RU_YAK) each grouped on their own.

3.2. Analysis of PCAdapt and BayeScan

The BayeScan analysis of the 30,708 SNP data set in 30 populations led to 1557 SNPs being identified as outliers. Of these outliers, 1349 SNPs were revelled to have a PO of 0.79876 or above with an FDR of no more than 5% and retaining the default prior odds of 10 (Figure 4) following the parameters as used in the study done in wolves by Schweizer *et al.*, (2016).

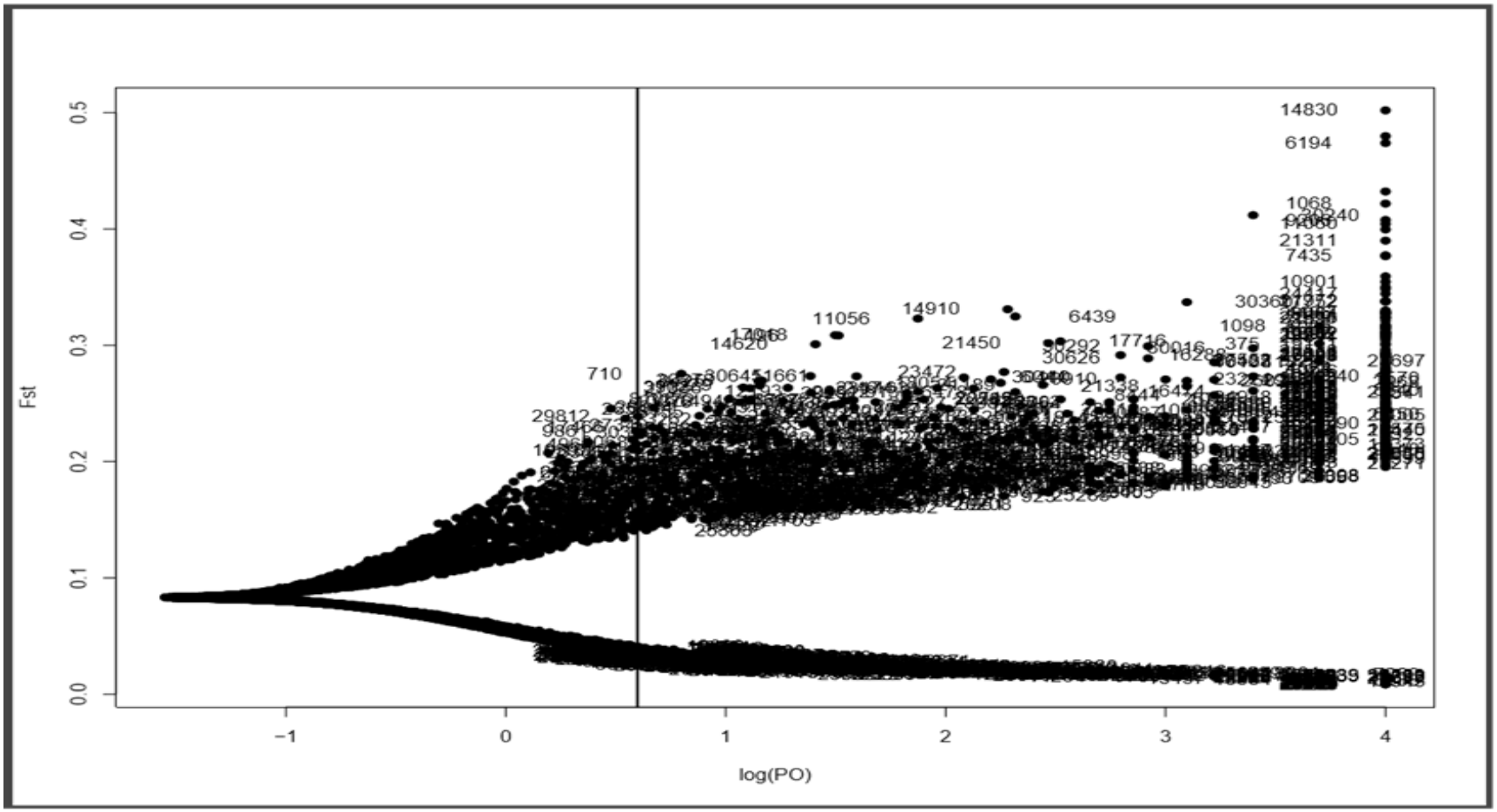

Figure *4* Results from a BayeScan method showing SNPs (black dots) on the left side of the PO threshold (black vertical line) suggesting no selection occurring. The SNPs on the right side of the PO line suggests selection has occurred, the SNPs group with the lower FST suggest balancing selection while the SNPs grouped with the higher FST suggest diversifying selection. Figure produced by using BayeScan (Foll & Gaggiotti, 2008).

From these SNPs, 794 (58.9%) were within a group with a higher FST (than the background FST) indicative of positive diversifying selection while the remainder fell in the group of lower FST (than the background FST) suggestive of balancing selection (Appendix I).

The PCAdapt analysis used resulted in 2,070 (outliers) SNPs using a K-value of three and with an associated q-value (p-value adjusted for the FDR) of 0.05. The range expansion method was used to take into account the admixture of the red foxes' ecological habits over its evolutionary history since the LGM (Statham et al., 2018). I also ran the same procedure with a K-value of two and four which produced 1,536 and 1,956 outliers respectively. Based on these results and the subjectivity of estimating the K-value (Zwick & Velicer, 1982), the decrease from K3 to K4, a K-value of three was used (Luu et al., 2017).

The overlapping outliers from these two methods with a single hit in BLASTn (161 SNPs) were located within 144 unique candidate genes. From this data 12 genes were found to be uncharacterised (Appendix II). All 161 of these SNPs were highly significant (P- value < 0.002) with the analysis PCAdapt implemented (Figure 5)

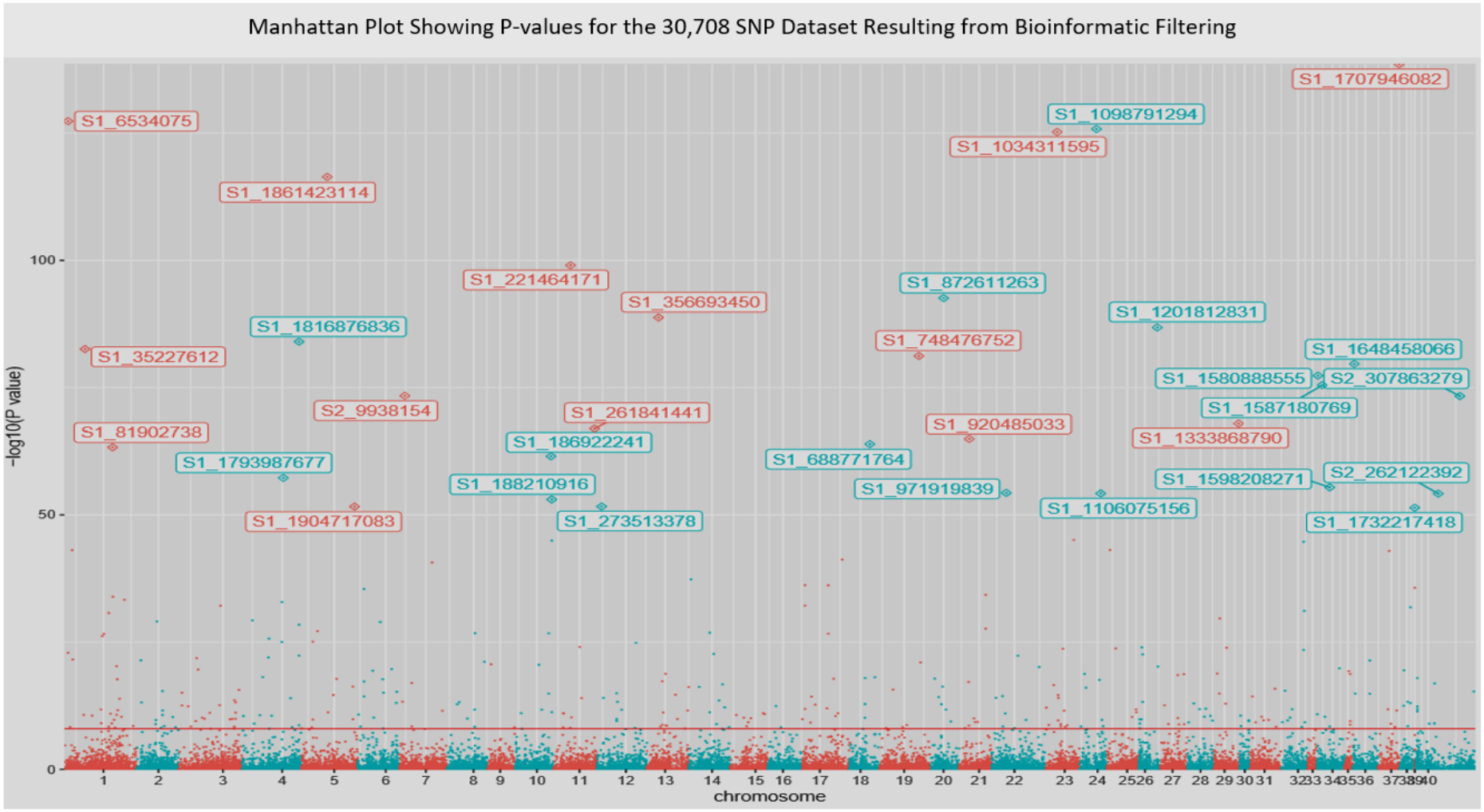

Figure 5 Shows a Manhattan plot from the 30,708 SNP data set. Significant p-values are shown above the red horizontal line and SNPs with a significance of 1e-50 or more are labelled. Chromosome 40 is representative of the X-chromosome and 39 are SNPs not assigned to a chromosome. Figure produced by using (RStudio Team., 2016).

The alpha and FST values generated through BayeScan ranged from 0.65 to 2.5 and 0.15 to 0.5 respectively (Appendix III)

3.3. Candidate genes

GO enrichment tests in g:Profiler placed 126 of these candidate genes into 62 GO category terms (Appendix IV and Figure 6). The GO category terms are divided in to molecular function (MF) terms (13), biological process (BP) terms (44) and cellular component (CC) terms (5) in g:Profiler with significant P-values (Figure 6 and Appendix IV).

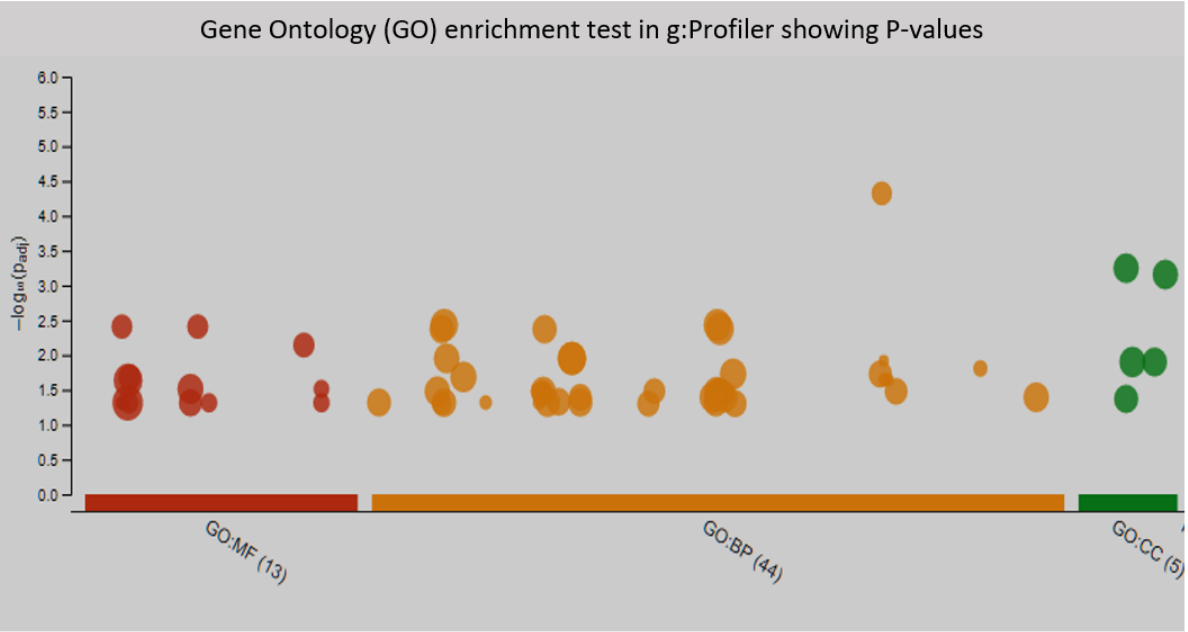

Figure 6. Displaying the results with P-values of the GO enrichment analysis using g;Profiler, the three broad parent term categories (MF, BP, CC) are displayed with child terms represented by the circles above. The circles are positioned with -log10 P-value (see Appendix IV for gene/s of each term). MF (Molecular Function), BP (Biological Process), CC (Cellular Component). Figure modified using https://biit.cs.ut.ee/gprofiler/gost.

In DAVID 142 genes were found relating to *C. lupus familiaris*. Using functional annotation clustering at an Ease-score of 0.05 in DAVID resulted in 38 clusters with 130 different terms (Appendix V and VI). Each cluster of terms have a similar biological meaning owing to sharing similar gene members (Huang, Sherman, & Lempicki, 2009c).

The 62 and 130 significantly enriched GO categories from g:Profiler and DAVID relating to subjects such as morphology (e.g. anatomical structure morphogenesis, animal organ development), movement (e.g. locomotion), response to stimulus (e.g. negative regulation of response to stimulus), behaviour (e.g. walking behaviour), physiology (e.g. blood circulation), metabolism (e.g. positive regulation of RNA metabolic process), muscle (e.g. muscle contraction) and much more was found (Appendix IV, V, and VI).

By using the 'Root' approach, genes within these terms were studied in more detail and the table in Appendix IX provides a summary of a study produced for each gene.

3.4. Allele Frequency Analysis

The allele frequencies without further study shows no clear patterns (Appendix VIII) with the possible exception of the MUC19 gene. The MUC19 gene seems to have a clear allele frequency change in the SNP S1_1220483264 at position 13229819 on chromosome 27 from a guanine (G) to a cytosine (C) in four Nordic populations (FN_N, SW_N, SW_S, NOR_NE) (Figure 1b and 7).

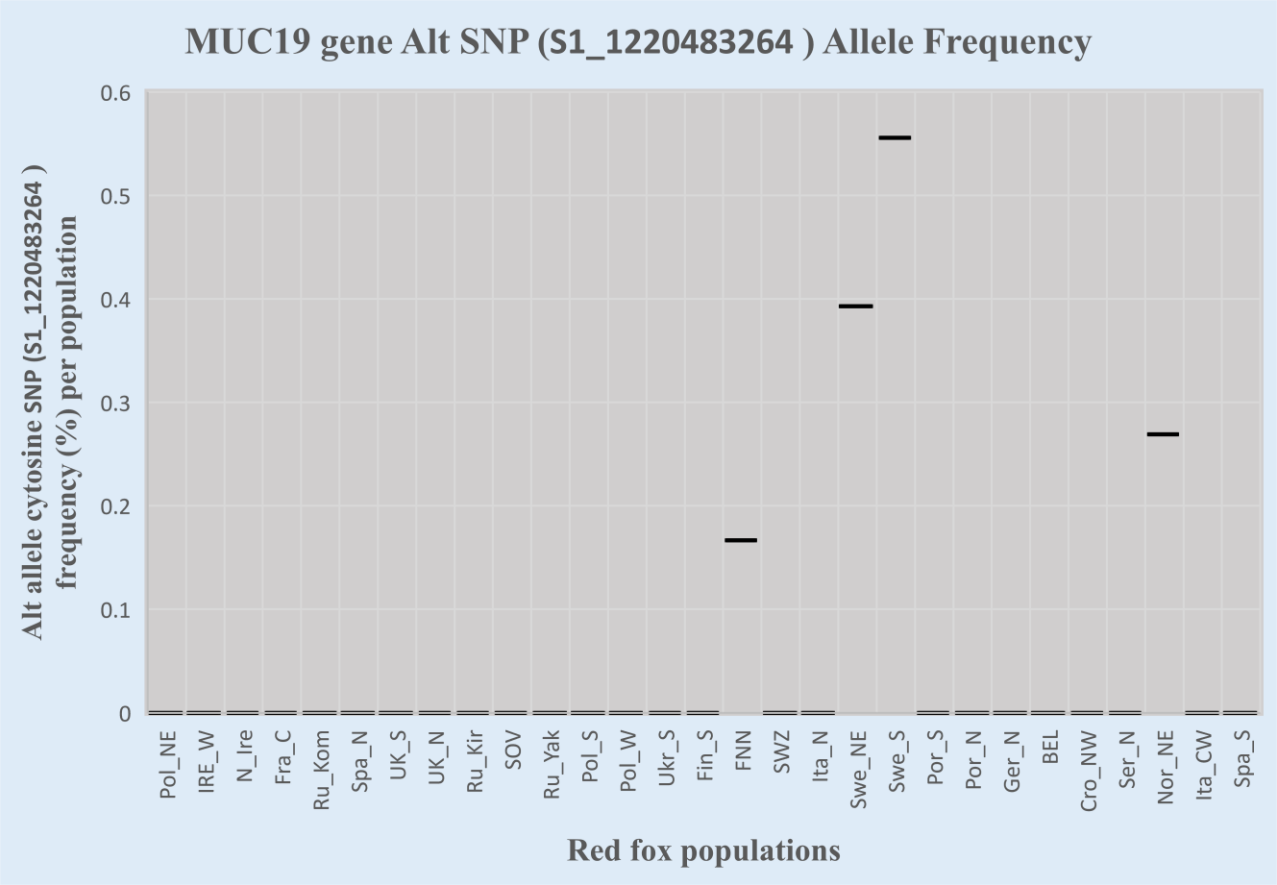

Figure 7. The populations Fin_S, Swe_NE, Swe_S, and Nor_NE show a distinct change in the allele frequency at position 13229819 on chromosome 27 from a guanine (G) to a cytosine (C). The remaining 26 fox populations have a guanine (G) fixation allele frequency.

4. Discussion

4.1. Population structure

Using a relatively large SNP data set , a clearer phylogeographic history for the red fox was revealed using a PCA-based analysis which was strengthened by having consistencies with previous studies (Edwards et al., 2012; Kutschera et al., 2013; Sommer & Nadachowski, 2006; Statham *et al.*, 2018). Additionally, my results along with previous studies (Edwards et al., 2012; Kutschera et al., 2013; Sommer & Nadachowski, 2006; Statham *et al.*, 2018) also seem to point to European red foxes occurring outside the traditional well-defined and separated refugia during the LGM. However unlike the study by Teacher, Thomas, & Barnes, (2011) I did find evidence of population structure to some degree.

I found the Iberian red fox populations to be isolated in a cluster on their own (Figure 3). These results indicate that the Iberian Peninsula red foxes have not interbred with close by populations for many generations which is consistent with mitochondrial evidence (Kutschera et al., 2013; Statham et al., 2018) and is possibly due to being isolated from other red fox populations by the Pyrenees mountains (Statham et al., 2018). This is considered one of the traditional refugia (Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998) and generally is an area of endemism rather than a source for northwards re-colonization of Europe in the postglacial period. This is corroborated by the results presented here.

Further north Edwards et al., (2012) found mitochondrial DNA evidence of Swedish red fox populations clustering with eastern European fox populations (and Denmark) while the Norway foxes were clustered with central European foxes. In addition, Statham et al., (2018) found a distant relationship between the Scandinavian populations and foxes from Russia. I found that the Scandinavian (Nordic) populations in a cluster close to the population groups of the Russian, central European, and Balkan countries' clusters. These results also seem to

indicate a recolonization in Scandinavia from red fox populations in Russia and Europe as the ice retreated which also follows the work of Statham et al., (2018). Additionally, I also found the central European and Balkan countries are tightly clustered around the Italian populations with the French and Belgium populations being more distinct in a very close-by cluster (Figure 3). This indicates admixture between the Balkan derived and Italian Peninsula derived populations, again following the work by Statham et al., (2018) and this inference is reinforced by fossil evidence linking these groups (Sommer & Nadachowski, 2006). Furthermore my results also suggest a past degree of gene flow existed between the Balkan populations (SOV, CRO, SER_N) and the populations in the east (URK and Russian populations) once more following Statham et al., (2018).

Further west, results from my PCA (Figure 3) shows the UK populations and Irish populations isolated from each other, and from most other European populations. The French and Belgian populations are found to be closest to the Irish (and with the Iberian populations) and the UK populations. This indicates that red foxes from the UK and Irish populations have been distinct for many generations, as supported by fossil evidence (Sommer & Nadachowski, 2006). Furthermore, Statham *et al.*, (2018) also indicates that the UK and Irish red fox populations were either founded by or was part of the French (and Netherlands) populations during and after the LGM. After the rise in sea levels the UK and Irish populations became isolated from the continent and the UK and Irish populations became isolated from one another.

Overall my results seem to suggest that although distinct red fox populations did exist in different refugia during the LGM, the extent to which these were separated populations is debatable. This conclusion was also found in previous studies (Edwards et al., 2012; Kutschera et al., 2013; Sommer & Nadachowski, 2006; Statham et al., 2018) and contrasts to the idea of red foxes like other taxa being completely isolated to well defined refugia during

the LGM (Taberlet et al., 1998) or existing as one continuous population over southern Europe (Teacher et al., 2011). This incomplete isolation of fox populations is most likely due to the dispersal capabilities and adaptable nature of the extremely plastic red fox.

4.2. Candidate genes

This study focused on diversifying selection and used methods which would reduce obtaining false positives while minimizing candidate genes not found. Most traits are polygenic from aspects of disease, morphology and red blood cell count (Shi, Kichaev, & Pasaniuc, 2016) to aspects of the immune system, learning, osteoporosis, and cancer (Matin & Nadeau, 2001). Therefore, not only can there be many genes involved for one trait, but these genes may also be involved with many traits themselves.

Consequently, it was important to pick only one (or two in some cases) functional aspect of each gene and focus in more depth on this. This knowledge was then used in relation to phenotypic traits of the red fox to begin to understand what possible selective pressures may have been occurring in which local adaptations arose increasing the fox's fitness. As this study found a large number of candidate genes it seemed necessary to assign them to larger groups or categories. A subset of the genes found in Appendix IX have been presented and divided into categories to produce a more detailed clearer study. Additionally, a few genes were included in more than one category if it seemed necessary

4.2.1. Immune system

The red fox, like other vertebrate's immune systems, is comprised of many different components which are categorized in two groups, the innate and acquired or adaptive

immune system. The innate immune system is composed of non-specific immune cells and physical barriers such as the skin which provides the first defence against pathogens while the more evolutionary recent acquired immune system includes B-cells and T-cells which target specific threats (Kimbrell & Beutler, 2001).

The red fox skin, like other mammalian skin, has many components that function to provide a permeable barrier against desiccation and a defensive barrier against pathogens (Elias, 2007). Of note, red foxes develop skin lesions when infected with some parasites such as the parasitic mite *S. scabiei*. Furthermore, variants in genes that are involved with barrier defence were found in this study. For example, by regulating genes involved with immunological responses, keratinocyte differentiation and epidermal development the LOC100688223 (FUCA1) candidate gene plays a role in the skin barrier defence system (Valero-Rubio, Jiménez, Fonseca, Payán-Gómez, & Laissue, 2018). By using a knock-down experiment Valero-Rubio *et al.,* (2018) found a change in expression of 387 genes of which 17 were found to be involved with keratinocyte differentiation/epidermal development and 61 genes involved with immune response. Valero-Rubio et al., (2018) also found that a down-regulation in LOC100688223/ FUCA1 is most probably involved with fucosidosis-related skin lesions.

The CERS3 candidate gene plays a key role in forming the epidermal permeability barrier which is vital to mammalian survival (Uchida, 2014). This epidermal permeability barrier may play an adaptive role in how much water is lost through the skin differing with foxes living in dry arid regions vs foxes living in a more humid habitat as it does with the house sparrow (*Passer domesticus L.*) (Cox, Muñoz-Garcia, Jurkowitz & Williams, 2008). Directly or indirectly ceramide also regulates the differentiation, proliferation and apoptosis in epidermal keratinocytes (Uchida, 2014). It has also been shown that mutations in CERS3 can

cause a disruption in the synthesis of ceramide leading to the formation of thickened, dry, scaly skin (ichthyosis) (Zaki & Choate, 2018).

Red foxes have many skin pathogens including some shared by humans such as the *Staphylococcus aureus* bacteria (Carson et al., 2012). PTPRK, like CERS3 and LOC100688223/ FUCA1, is another gene found in this study that plays a role in regulating the proliferation of keratinocytes. Keratinocytes among other functions produces antimicrobial peptides (AMPs) which kill microbes such as bacteria (Erin Chen, Fischbach, & Belkaid, 2018; Xu, Xue, Zhou, Voorhees, & Fisher, 2015).

Red foxes can also be infected by many other types of pathogens, if for example the skin is breached, such as viruses (Bourhy et al., 1999b), bacteria (Scholz et al., 2009), fungi (Malmasi, Khosravi, Selk Ghaffari, & Shojaee Tabrizi, 2009), and helminths (Reperant et al., 2007). Candidate genes found in this study relating to infection include the following, RIPK1, RTP4, RAB30, WDR41, NLRP8, SELP, and TACR1. During an infection from bacteria the RIPK1 candidate gene plays a vital role in activating the immune system. This occurs by inducing a kinase–induced apoptosis which in turn promotes antibacterial defence and innate immune cytokine production (Peterson et al., 2017). Whereas, during a viral (e.g. Japanese encephalitis virus) infection the RTP4 gene is upregulated (Gupta, Santhosh, Babu, Parida, & Rao, 2010). Red foxes have been found to be infected by viruses such as the rabies virus (see introduction), common bacteria such as *S. aureus* (Carson et al., 2012), and more recently even a new bacterial infection from *Brucella microti* has been found (Scholz et al., 2009). These genes may play a role in red foxes' immune system and evolve to suit specific fox populations needs depending on the types of virus or bacteria found in each region.

Other genes that may play a more general role red fox immunity include the following. The RAB30 gene may play a role in immunological processes by regulating autophagy against

pathogens (Oda et al., 2016) while WDR41 is also involved in the autophagy-lysosome pathway (Hu et al., 2016). Finally, when the TACR1(NK1R) gene product interacts with the substance P ligand the activation of nuclear factor-kappa B (NF-κB) and proinflammatory cytokines occurs. The signals produced by the proinflammatory response are very important in fighting off parasitic diseases and other pathogen infections (bacterial, viral, and fungal) as well as the general functioning of the immune system (Douglas & Leeman, 2011). The red fox's evolution has been influenced by parasites such as the rabies virus and the scabies mite (*S. scabiei*) as discussed in the introduction and so variants in these genes is understandable.

Most noteworthy, this study found that the candidate gene MUC19 reached fixation (1) for the major allele in all European populations except in four Nordic populations (Figure 7). MUC19 is a member of the mucin family. Mucins form a protective layer of mucus which provides a protective barrier against inhaled particles and pathogens which are trapped and subsequently actively transported out of the respiratory tract by a process called mucociliary clearance (MCC). The mucus produced contains antiprotease and anti-microbial and other detoxifying properties and so acts as part of the innate immune defence system (Weitnauer, Mijošek, & Dalpke, 2016).

Interestingly Forchhammer & Asferg, (2000) highlights several studies that documents the spread of Sarcoptic mange disease throughout most of the Nordic red fox populations during the 1970s and 1980s. This disease reduced the Nordic red fox populations by over 70%. The conditions to which preceded these fox populations crashes in a relatively small time period may have been due to the rapid increase in density and number of red foxes which urban living provides (Goszczyński, 2002). This may have increased conspecific interactions between red foxes and even allowed new diseases to adapt to these new (relatively crowded) conditions of high fox abundances (Goszczyński, 2002). Furthermore, complex interactions

between macroparasites and microparatites can have a negative cross-regulatory effect (immunosuppressive) on a vertebrates immune system (Jolles, Ezenwa, Etienne, Turner, & Olff, 2008) and disease can be a major evolutionary selective force. Therefore, the MUC19 correlation in these four fox populations may be showing a change in the MUC19 frequencies due to a changed selective pressure having occurred in that area.

This evolutionary arms race between host immune system and parasite is a major selective driving force in evolution which has occurred since the red fox species first evolved. Some pathogens are more prevalent in some areas, some have distinct geographical separations and some overlap only slightly such as *U. stenocephala* occurring in more vegetated areas by the coast and *D. caninum* occurring in inland areas where there is less vegetation (Dybing, Fleming, & Adams, 2013a). These two species of parasites (*U. stenocephala* and *D. caninum*) infect red foxes and have distinct distributions for the most part (Dybing, Fleming, & Adams, 2013b). A fox population diverging from one habitat (e.g. costal) to another (e.g. inland) may bring the new population of foxes in to contact with other previously unencountered species of parasite. These newly encountered parasites may have restricted distributions, and this may impact what immune system variants will be most advantageous in each fox population across Europe. However, some pathogens may have a large distribution.

A relatively common pathogen that infects the red fox (intermediate host) is *T. gondii* which is a coccidian protozoan and is globally distributed. The definitive host of *T. gondii* is felids, the pathogens oocysts are shed in the cat's faeces and then can spread to many different mammal and bird species (eating these animals is how the foxes get infected) (Dubey, 1998). Prestrud et al., (2007) references several studies done on the prevalence of *T. gondii* in red foxes (foxes with antibodies against *T. gondii*) in many European countries to be 20–38% and up to 68% in Hungary. This is noteworthy as my study found the candidate gene NLRP8 and

the expression of NLRP8 mRNA was found to be upregulated during infection with *T. gondii* which results in inflammation activation (Chu et al., 2016).

T. gondii forms cysts in its host, often in the brain, which can result in a chronic infection. In this chronic infection there seems to be an evolutionary stability between the host's immune defence and the parasite's evasion of the immune response (Carruthers and Suzuki, 2007). *T. gondii* can also be transmitted to unborn young from the mother in some species such as the arctic fox (*Alopex lagopus*) (Prestrud et al., 2007) and so from one generation to the next.

Furthermore, selective pressures exerted on red foxes caused by pathogens may be further reaching. It has already been discussed above that *T. gondii* cysts are found in the brain (Carruthers & Suzuki, 2007) and can be transmitted to the next generation (Prestrud et al., 2007). Candidate genes found in this study are associated with neurological disorders and dogs with *T. gondii* show neurological deficits (Lempp et al., 2017). Finally, behavioural traits associated with this infection in mice and perhaps humans has been suggested (Webster, Lamberton, Donnelly, & Torrey, 2006).

4.2.2. Nervous system: behaviour, learning and memory

The mammalian nervous system is complex with many different specialized cell types interacting which results in the complexities of learning, memory and behavioural actions (Zeisel et al., 2018) (amongst many other functions) seen in organisms such as the red fox.

Many candidate genes relating to behavioural phenotypic traits were found in this study including the following, NBEA, PLCB1, TACR1(NK1), TENM2, CTNNA2, and SORCS2. This is not surprising as the red fox is very adaptable. The NBEA (Neurobeachin) candidate gene I found produces a scaffolding protein in the brain and is involved with synaptic functioning and neurotransmitter release. Studies with mice with one allele of NBEA deleted

(haploinsufficiency) had dramatic phenotypic effects such as a decrease in sociability, an increase in self-grooming, a heighted conditioned fear response to stimuli, and most notably changes in the hippocampus and impairment in spatial learning and memory (Nuytens et al., 2013). Red foxes often live in social groups which may include helpers of a dominant pair. In times of stress such as a high population density (which can especially be found in urban areas) the dominant pair of red foxes may suppress the reproduction of female helpers (Macdonald, 1979). This stress of suppression in rodents has been shown to alter hippocampal neurogenesis (Gheusi, Ortega-Perez, Murray & Lledo, 2009).

Like NBEA, other genes involved with spatial learning and memory was also found. The PLCB1 gene encodes the PI-PLC β1 enzyme. Spatial memory deficits, working memory and other behavioural abnormalities occurred in Pi-plc β1 KO mice (Vasco, Cardinale, & Polonia, 2012). It is not surprising this study has found several candidate genes which play a role in spatial learning and memory as foxes are known to show great skill in caching and retrieving food caches (sometimes this is done to feed their young) (Macdonald, 1979; Macdonald, Brown, Yerli & Canbolat, 1994).

Other behaviour candidate genes such as the TACR1(NK1) encodes Neurokinin 1 receptors (NK1R). These receptors, along with the neuropeptide substance P (SP), are richly distributed in the mammalian nervous system including structures in the brain's fear network such as the amygdala (Hoppe et al., 2018a). The SP-NK1 system has been shown to regulate stress and anxiety-related behaviour in previous studies (Bilkei-Gorzo and Zimmer, 2005; Ebner and Singewald, 2006; Hoppe *et al.*, 2018). Hoppe *et al.*, (2018) concludes their study by demonstrating a positive correlation between amygdala NK1 receptor availability and trait anxiety and negatively with extraversion. TENM2 is also a gene that is primarily expressed in the brain where it is involved with regulating synaptic connections (Tews et al., 2017).

Encoding for a cell-adhesion protein called alpha N-catenin it is thought CTNNA2 is important in regulating synaptic plasticity and a knockout homologous gene (Catna2) in mice produces defects in the brain and its functions correlating to impaired response in fear conditioning, enhanced acoustic startle responses and cerebellar ataxias (Ehlers et al., 2016). Behavioural differences in red foxes have been documented and may be in part due to variants in these behavioural related genes. For example, red foxes are known to be more curious and exploratory in their habits while in zoo conditions compared to other carnivores (Amrein & Slomianka, 2010). When living in urban areas however red foxes remain vigilant of humans and are shy (Amrein & Slomianka, 2010). Another gene, the IL1RAPL1 mediates excitatory actions in neurons. In IL1RAPL1 knock outs, mice increased their locomotor activity and impairments of cognitive functions in several learning tests were documented (Yasumura et al., 2015). IL1RAPL1 is also involved with brain functions including learning, memory, behavioural flexibility, locomotor activity and anxiety (Yasumura et al., 2015). Variants in genes playing a role in the brain which may contribute to behavioural changes in red foxes may be due to human disturbance selective pressures (among others) which may also result in foxes being nocturnal in urban areas to avoid human conflict (Díaz-Ruiz et al., 2016b).

The hippocampus region of the brain is also enriched with the SORCS2 protein and also has effects on synaptic plasticity and spatial learning (Glerup et al., 2016). SORCS2 knockout mice are associated with bipolar disorder, schizophrenia, and attention deficit hyperactivity disorder (ADHD) (O'Rourke & Boeckx, 2018).

Interestingly a study by Kukekova *et al.*, (2018) found strongly different behavioural phenotypes in a long term selective breeding programme in a Russian farm-fox experiment. By selecting a subset of red foxes from the tame and aggressive individuals from a common

population and analysing their genomes they found regions with significantly decreased heterozygosity and increased divergence between the aggressive and tame populations. A strong candidate gene for the tame behavioural trait in red foxes was SorCS1. This gene seems to play a role in synaptic plasticity in fox domestication (Kukekova et al., 2018).

In my study I found two SNPs in the candidate SorCS2 gene which is in the same family (Vps10p-domain receptor family) as the SorCS1 gene (Hermey, 2009) found by Kukekova *et al.*, (2018). These two genes (SorCS1 and SorCS2) along with some other candidate genes mentioned above seem to play a role in synaptic plasticity (Glerup *et al.*, 2016; Kukekova *et al.*, 2018) and behaviour. Both SorCS1 and SorCS2 are highly expressed in the mammalian brain (Hermey, 2009) and so it is conceivable that the SorCS2 candidate gene found in my study is also involved with behavioural traits which may be selected for in wild red foxes in nature.

In nature red foxes may have very large home ranges and may disperse large distances (see introduction) so genes relating to behaviour, spatial learning and other learning skills may come under selective pressures. The candidate gene found called PTPRD (protein tyrosine phosphatase gene) is highly expressed in the brain where it plays an important role in organizing excitatory and inhibitory synapses. It has been shown with PTPRD knockout mice that this gene is important for memory and learning in spatial tasks and early mortality occurred due to reduced food intake. The FOXP2 is expressed in brain regions that are known to be important for motor-related functions and in motor-skill learning (French et al., 2012) which may be important depending on the prey of the particular population of red foxes. For example, rural fox populations tend to have more rodents in their diet than urban foxes and so may need more refined reflexes to obtain adequate food resources. The FOXP2 is also known to be important in the development and use of vocalizations in humans and other mammals to

communicate (Nielsen et al., 2007) and red foxes may communicate vocally over short and long distances depending on their habitat. ANO2 is implemented in learning plasticity, behavioural experiments with knockout mice has also revealed the positive role that ANO2 also has in motor learning. Mice with the knockout ANO2 and/or KCNMA1 gene have reduced motor learning abilities and showed mild ataxia (Bentzen *et al.*, 2014; Neureither *et al.*, 2017). Variants in these genes may be important to red foxes as they catch prey, cache food, and retrieve it at a later date (as mentioned above) (Macdonald, 1979; Macdonald, Brown, Yerli & Canbolat, 1994). This implies that motor learning, spatial awareness, and memory is important in foxes' lives and variants may increase fitness in some populations but may have high costs in other populations.

Other candidate genes found which may be important to red fox skills are the PTPRD and DLG2. PTPRD knockout mice also shown poor posture, growth retardation, and motor skill defects (Choucair et al., 2015). Finally, DLG2 (another candidate gene found) has an important role in brain plasticity, synapses development, stability, complex cognitive tasks and learning tasks (Perrone et al., 2017).

The importance of being able to construct a cognitive map to an animals surrounding environment was first hypothesized by Tolman, (1948). Learning the spatial relationships between different areas to feed, to den, territory boundaries, and a host of other important places and importing these into memory seems to be likely in red foxes (Fabrigoule and Maurel, 1982). For example, as red foxes pushed into higher latitudes with the receding ice age there may have been less food and so red foxes would require larger home ranges (Goszczyński, 2002; Huang *et al.*, 2017). More recently red fox populations living in urban environments have access to a large food resource and so need smaller territories (Goszczyński, 2002; Huang *et al.*, 2017) but have to adapt to more competition from

conspecifics. This difference in territory or home range size as well as living with a high abundance of other foxes may have an impact on variants of the genes above involved with spatial learning, behavioural flexibility, locomotor activity and a genetic tendency to explore.

Furthermore, urban areas also have roads which present a significant risk to wildlife and may restrict gene flow if they cannot be crossed. This may have the effect of isolating fox populations from each other (Wandeler et al., 2003) resulting in allele frequency changes in different genes in these isolated populations. These isolated populations in urban areas have less genetic diversity than their rural counterparts (Wandeler et al., 2003) which may be due to adapting to a similar environment with similar selective pressures or perhaps less likely due to a founder effect. Mortality on roads due to traffic in red foxes may be a significant selective pressure by the way of young foxes that are less cautious being taken out of the gene pool before reproducing. This would leave the older foxes to have developed behavioural skills to avoid death by judging speeds of cars and crossing at safer points and times (Baker, Dowding, Molony, White, & Harris, 2007). Learning to be road safe is an essential skill for urban foxes especially during the dispersal period. Additionally, it has also been documented that roads are sometimes used as preferred pathways on their movements in their territories (Fabrigoule & Maurel, 1982). Therefore, it seems plausible a divergence in populations led to different challenges and a change in allele frequencies for learning, memory, and behavioural related genes. However, further investigation would help to rule out or strengthen proposed hypotheses.

Living close to human settlements, especially where the culling of foxes occurs, and in areas where predators of red foxes occur, seems to have the effect of foxes restricting their active time to when there is more darkness (Díaz-Ruiz, Caro, Delibes-Mateos, Arroyo, & Ferreras, 2016a). However if there is dense vegetation even in areas disturbed by humans foxes are

active in the daytime too (Díaz-Ruiz, Caro, Delibes-Mateos, Arroyo, & Ferreras, 2016b). This study suggests that habitat type, human disturbance and predators also have influences on red fox behaviour and possibly on sensory adaptations.

4.2.3. Nervous system and senses

Red foxes are mainly nocturnal and crepuscular animals which is influenced by the presence of humans, prey availability, habitat type and season (Díaz-Ruiz et al., 2016b). During the dark, red foxes rely predominantly on hearing to capture rodent prey (Wells & Lehner, 1978). Several genes involved with hearing was found in this study such as the gene called protocadherin related 15 (PCDH15). This gene is predominantly expressed in the Central Nervous System (CNS). The expression of PCDH15 is highly expressed in the sensory epithelium in the developing inner ear and the neural and pigment layers of the retina and lens capsule epithelium (Murcia & Woychik, 2001). Mutations in PCDH15 have been shown to cause nonsyndromic deafness (DFNB23) or impairment of hearing and vision in humans (type 1 Usher syndrome; USH1F) (Ahmed et al., 2008). Lack of expression in the PTPRD gene and mutations in EYA1 gene have also been implicated in hearing loss (Choucair *et al.*, 2015; Namba *et al.*, 2001).

Red foxes can hear a lower frequency (51 Hz) than domestic dogs (67 Hz) or domestic cats (55 Hz) and a higher frequency (48 kHz) than the domestic dog (45 kHz) (Malkemper, Topinka, & Burda, 2015). Hearing is very important to red foxes and the expression of the SULF1 candidate gene found is important in the development of the avian and mammalian ear. SULF1 plays a role in the development of mechanosensory hair cells which are vital for the perception of sound (Freeman, Keino-Masu, Masu, & Ladher, 2015). While Kcnma1 channels in hair cells allow the detection of high-fidelity, high precision sound processing.

The KCNMA1 candidate gene found has many functions of which the Kcnma1 channels production is one (Molina et al., 2013).

The KCNMA1 candidate gene is also important with eye blink reflex (Yeşil et al., 2018) and the EYS gene is important in the development of the eye and in the photoreceptor cells of the retina where the EYS gene is highly expressed (Hashmi et al., 2018). The visual system of red foxes is thought to be like dogs and adapted to function in aspects such as in various light conditions and motion detection (Pongrácz, Ujvári, Faragó, Miklósi, & Péter, 2017). Interestingly, a couple of candidate genes were found with variants relating to these traits. The ARRB1 candidate gene plays a role in phototransduction of rod cells which is important for scotopic (low-light) vision (Lamb, Patel, Chuah, & Hunt, 2018) while MYH13 is expressed in extraocular muscles which functions to move the eye (Briggs & Schachat, 2002).

Other genes found to be involved with the processes and mechanisms of the sensory perception of sight include DSCAML1, TSHZ2, and SCFD2. DSCAML1 is involved with retinal and other neurodevelopmental processes (Fuerst et al., 2009), TSHZ2 is expressed in hindbrain and the neural retina during development (Santos, Fonseca, Vieira, Vieira, & Casares, 2010) and SCFD2 is thought to play a transporter role with proteins in the ocular tissue (Ricketts, Pettitt, McLaughlin, Jenkins, & Mellersh, 2015).

Red foxes inhabit a wide range of environments from closed temperate forests where light conditions vary only slightly to the arctic tundra where light conditions vary to extremes with the seasons. Hearing may be more important over much of its range where it is active in low light conditions, in the far north where winters are long and dark, where there is snow and where prey items are out of visibility (Červený, Begall, Koubek, Nováková, & Burda, 2011). However, although the selective pressure on hearing sensitivity may be thought to have

evolved primarily to detect mammalian prey such as rodents moving in grass or under snow, Malkemper, Topinka and Burda, (2015) argues it is more likely to have evolved to detect avian prey over larger distances. Hearing may also be important in forest habitats as vegetation could reduce visibility of prey as it may do for the coyote (Richer et al., 2016).

However, in forest habitats red foxes may find carcasses to be an important food source (Jędrzejewski & Jędrzejewska, 1992). Furthermore, high altitude mountain habitats such as in the Sˇumava mountains of Croatia (Bohemian Forest) red foxes may rely more on smell to find carcasses that have been killed by other predators (Hartová-Nentvichová, Šálek, Červený, & Koubek, 2010).

The ANO2 is a candidate gene found in this study related to olfaction functions and it belongs to the Anoctamin (Tmem16) family of calcium-activated chloride channel genes. This gene is expressed in the main olfaction epithelium, olfactory sensory neurons, vomeronasal organ, and the olfactory bulb. It has been shown with ANO2 knockout mice studies that although not essential for olfaction this gene may boost the sensitivity of olfaction giving a slight evolutionary advantage over animals that have a non-functional Ano2 gene. This goes against the classic view that Ano2 and other calcium-activated chloride channel genes may have a substantial role in mammalian olfaction (Billig, Pál, Fidzinski, & Jentsch, 2011). Another gene involved with cilia called CFAP74 is suspected to have a role in olfaction, however little is known about this genes function (Dong et al., 2017). Changes in these olfaction related genes may be more important for foxes in high altitude red fox populations relative to lowland populations.

Higher latitudes also usually have more snow, this affects the hunting success of red foxes and other predators which may rely more on olfaction and auditory senses when hunting small rodents under these conditions (Halpin & Bissonette, 2008). While further investigation

is needed, the varying light conditions, the availability of food detected by smell, and snow and vegetation density may have exerted diversifying selection pressures on red fox senses over the diverse habitats and latitudes it lives which may also affect fox morphology.

4.2.4. Bone, skeletal muscle, limbs, and dentition

ADAM12 has been found to be expressed in various tissues (e.g. mesenchymal cells surrounding the intestine and lung, ensheathing cells of peripheral nervous system in mouse embryos, and brown adipose tissue) most prominently here is that it is known to be important in the development of muscle and its regeneration (Kronqvist et al., 2002). Like other ADAM genes ADAM12 is highly conserved between teleost fish and mammals (Tokumasu et al., 2016). By using ADAM12 knock-out zebrafish Tokumasu *et al.*, (2016) found a reduction in body size compared to the wild type specimens although morphologically no other difference. This may be due to adam12 being involved with regulating genes affecting cartilage/bone development and so growth (Tokumasu et al., 2016).

Changes in body size would have also show effects on other genes too, I found candidate genes relating to bone, limb and tooth development. GPC6 is involved directly with bone formation with studies on mice displaying significantly shorter long bones than would normally occur if both GPC6 alleles were muted (Capurro et al., 2017). Whereas, Sema6D is vital for osteoclastogenesis (bone moulding cell generation) which in turn is critical for the correct development and maintenance of the skeleton (Kang & Kumanogoh, 2013). EDIL3 is also involved with osteoblast differentiation which occurs as products from EDIL3 trigger integrin α5β1 signalling-mediated ERK-dependent Runx2 expression. This in turn up regulates extracellular matrix (ECM) genes such as alkaline phosphatase (ALP) and osteocalcin (OC) which regulates mineralization during bone development and remodelling

(Oh et al., 2017). Finally EDIL3 also seems to be important for correct development of the bones of the skull and face (craniofacial development) (Oh et al., 2017).

Changes in body and craniofacial sizes may impact tooth structure and development and candidate genes found in this study include EDA and FAM155A. EDA is involved with guiding the number and shape of teeth that develop (Tucker & Sharpe, 2004) while FAM155A plays an important role in the number of teeth (supernumerary teeth derived stem cells (SNTSCs)), in their proliferation and migration, unrelated to age (Lu et al., 2019).

Muray and Larivière, (2002) found (after considering sex and age) a significant difference in foot area size between lower and higher latitude red fox populations (n= 199). The higher the latitude the larger the size of the foot area with red fox populations found north of 48 degrees-N being 12% larger than further south. Genes found relating to limb development include the following; PTCH1 has been shown to interact with Shh resulting in limb morphogenesis such as autopod asymmetry and digit reduction in cattle (Lopez-Rios *et al.*, 2014; Ozretic *et al.*, 2016), AFF3 is expressed in the mouse developing limb including in the subectodermal mesoderm and particularly in the dorsal and ventral mesoderm (Feenstra et al., 2012), and the mouse ortholog (2210408I21Rik) of KIAA0825 is expressed in all critical points during the developing limb processes and it is thought it may also be involved with digit formation (Ullah et al., 2019).

Sablin and Germonpre, (2004) found that, apart for the most northerly red fox populations (probably due to lack of food resources), Bergmann's Rule is followed by the red fox in that with increasing latitude the size of the red fox gets larger. This is consistent with findings from Muray and Larivière, (2002) with larger foot area sizes with increasing latitude. Red foxes were confined to lower latitudes during the LGM (Statham et al., 2018) and moved north as the climate got warmer this would suggest a fluctuation in body size would occur. In

this study many candidate genes involving bone, tooth, and size development have been found of which some have been discussed above. These changes could be due to red foxes expanding their range north in which case there should be a distinctive shift in allele frequencies between those in the higher and those in the lower latitudes. However, like latitude, altitude also has an effect on body size as Bergman's rule also applies to temperatures as shown in the study done by Sablin and Germonpre, (2004). So temperature, latitudinal and altitudinal effects would simultaneously have to be considered as well as food abundance (Sablin & Germonpre, 2004) when looking for allelic frequency associations in genes relating to morphological, metabolic, and physiological adaptations.

4.2.5. Physiology

Red foxes can adapt to local conditions such as in different altitudes which is shown by the montane foxes living in Sacramento Valley (Sacks et al., 2010). Swanson, Fuhrmann and Crabtree, (2005) also shows distinct genetic differentiation between high altitude (>2300 m), mid-altitude (1600–2300 m), and low-altitude (<1600 m) red fox populations which is possibly due to the lack of gene flow to these isolated populations. My study has found candidate genes relating to high altitude adaptations in European red foxes.

The candidate gene Zfp90 found in this study has a role in many diverse biological functions by regulating gene expression. Zfp90 is vital for the proliferation and differentiation of hematopoietic stem cells (HSCs) (Liu *et al.*, 2018). Two SNPs were found in the Zfp90 gene which may suggest this gene has been important in red fox evolution in either maintaining HSCs or some other function. Another candidate gene found that is also expressed in HSCs and is important to stem cell self-renewal and hemopoiesis is MECOM (Zhang et al., 2011).

Hematopoietic stem cells are produced in the bone marrow where they differentiate into erythrocytes (red blood cells), leukocytes (white blood cells), and platelets (Mackey, 2001). The leukocytes are involved with the immune system and so with different diseases in different areas of the red foxes range disease may be driving diversifying selection in HSC related genes. However, as red foxes spread from their refugia they would have encountered habitats which would have differed in altitude. As they settled in these new areas their erythrocyte abundance may have changed (Peng Li et al., 2011) as individuals having a genetic variation which resulted in an advantageous phenotypic erythrocyte count at higher altitudes may have resulted in a higher fitness.

Higher red blood cell (RBC) counts in Tibetan chickens (*Gallus gallus domesticus*) reared at a higher altitude were similar to the high altitude native white-tailed ptarmigan (*Lagopus leucura*) RBC counts. Additionally, this study also found a lower RBC count in Tibetan chickens reared at a lower altitude (Zhang, Wu, Chamba & Ling, 2007). Likewise, mammals seem to also show higher RBC counts in populations native to a higher altitude (compared to populations living at a lower altitude) (Bunn & Poyton, 1996; Wu et al., 2005).

Red fox populations living at high altitudes such as in the Carpathian region (Malczewski et al., 2008) may become genetically isolated as red fox populations in the Sacramento Valley (Sacks et al., 2010) and the Greater Yellowstone Ecosystem in USA seem to have done so (Swanson et al., 2005). This isolation may cause a change in allele frequency in different altitudes in which red foxes live. This may have caused diversifying selection to occur in high altitude physiological related genes such as the Zfp90 and MECOM genes and a phenotypic RBC count change to occur.

Zfp90 also plays a role in the regulation of the developing heart (Liu *et al.*, 2018) as do the SEMA6D (Toyofuku et al., 2004) and TNNI3K (Wang et al., 2017) candidate genes also

found. The Neuron-restrictive silencer factor (NRSF), which itself binds to specific sequences of DNA to down regulate genes downstream in many different biological processes including cardiovascular homeostasis. Zfp90 binds to NRSF and prevents it from binding with the DNA. During pathological stress Zfp90 expression is up regulated resulting in the reactivation of the normally (after birth) quiescent foetal cardiac gene program which leads to cardiac remodelling (Hata et al., 2011). TNNI3K also seems to play a role in cardiac remodelling, Vagnozzi *et al.*, (2013) found that inhibiting TNNI3K limits chronic adverse remodelling while maintaining cardiac function. Interestingly, Stembridge *et al.*, (2014) shows that structural and functional remodelling of the Himalayan native Sherpa heart has occurred, and this may be due to genetic adaptation. The expression of another candidate gene, GALNT13, is thought to be regulated by hypoxia and may be involved with pulmonary vascular remodelling (Desai et al., 2013).

Another interesting point regards a SNP that was found within the RYR2 gene which has also been found to be a hypoxia-related gene (Zhang et al., 2014). RYR2 enhances Ca2+ release resulting in the contraction of the heart (Diviani, Maric, Pérez López, Cavin, & del Vescovo, 2013). This candidate gene was also identified as a positively selected hypoxia-related gene in the Tibetan grey wolf (*C. lupus chanco*) living in a high altitudinal environment (>3000) (Zhang et al., 2014).

Finally, by using exon boundary-spanning primers and Polymerase Chain Reaction (PCR) Cunha *et al.*, (2008) was able to show that the ANK2 gene with its 53 exons can produce a multitude of alternative splice variant polypeptides. The subsequent consequence is multifaceted ankyrin-B gene products which can specialize in different functional aspects in cardiomyocytes and cardiovascular physiology. One of these aspects may prevent genetic cardiac arrhythmia during hypoxic conditions (Li & Zhao, 2009).

Genes that may be selected for in foxes living at a high altitude may include genes that are also related to cold adaptation as high altitudes and cold environments usually occur together.

4.2.6. Metabolism and thermogenesis

Careau, Morand-Ferron and Thomas, (2007) describes how the red fox's basal metabolism is affected by climate. The adaptations the red fox may have to an arctic, cold or high-altitude environment can include larger organs (relative to individuals in a warmer environment) which can help boost their thermogenic capacity (Careau et al., 2007).

The PPARGC1A (PPARG coactivator 1 alpha) is a thermogenic gene involved in the Development of brown/beige adipocytes. The candidate gene PPARGC1A found in this study may have been targeted by natural selection to produce a cold-adaptive variant during the last glacial period in red foxes. TRIB2 (a gene not found in this study) is a thermogenic gene similar to our candidate PPARGC1A gene and TRIB2 variants are thought to be an adaptation to the climate of the ice age in some animals (Nakayama & Iwamoto, 2017) which may include the red fox.

PPARGC1A also interacts with another candidate gene found (FOXO1) to regulate the metabolic pathway gluconeogenesis (Puigserver et al., 2003) which may help in areas where there may be little food such as snowy mountain habitats and in the higher latitudes of the red foxes range. While FOXO1 also plays a vital role in adipocyte differentiation (Munekata & Sakamoto, 2009).

The THRB (thyroid hormone receptor beta) candidate gene in mammals is also involved with adaptive thermogenesis by regulating the expression of a mitochondrial uncoupling protein (UCP-1) that enables the generation of heat in brown adipose tissue (Lowell & Spiegelman,

2000; Tigano, Reiertsen, Walters & Friesen 2018). TENM2 is also involved with the regulation of UCP1 and hence brown adipose tissue. Interestingly by using Small Interfering RNA (siRNA) to remove TENM2 (gene knockdown) from preadipocytes it was shown to have the effect of increasing UCP1 and so increase mitochondrial respiration. TENM2 maintains white adipocyte tissue where it is highly expressed (Tews et al., 2017).

It could be that these genes involved with thermogenesis and adipocyte differentiation may be showing diversifying selection as the climate warmed during the end of the most recent ice age. However cold areas to the north and at higher altitudes would remain in which red foxes could have moved into. This study has found candidate genes involved with regulating UCP1 and interestingly Nishimura *et al.*, (2017) hypothesizes that UCP1 gene was important to another large mammal's (humans) ability to adapt to cold areas.

By looking at the allele frequencies in theses thermogenic related genes in this study, a correlation should be found between those populations living in cold environments and those populations living in warm environments. Since the red fox can live at temperatures down to minus 13 degrees Celsius before having to use thermogenesis (Careau et al., 2007) this seems like a good temperature to use when looking at the different populations for these cold adaption related candidate genes.

Of course, red foxes do live in cold environments and so the genetic variation from the foxes that best suited these habitats must have been passed to the next generation. However, climate, food abundance, and latitude may also affect adaptations in reproductive traits in foxes and other canids (Asa and Valdespino, 1998).

4.2.7. Reproduction

Red foxes taken from one hemisphere and relocated to the other seem to shift their breeding cycle by six months which points to photoperiodic cues being involved in regulating the breeding cycle in these seasonal breeders (Asa and Valdespino, 1998). Martins *et al.*, (2006) also points to a study that found maximum fertility in red foxes that was found to occur with increasing day length at high latitudes (>60°N) and at lower latitudes (<50°N). This suggests it is a decreasing day length that is associated with fertility in red foxes (and other canids). Furthermore, the testicular size and spermatogenic activity of male blue foxes (*Alopex lagopus*) changes significantly depending on the season (Smith, Mondain-Monval, MOller, Scholler, & Hansson, 1985). Environmental cues can cause hypothalamic release of hormones which regulate gonadotropins, such as follicle stimulating hormone (FSH) and luteinising hormone (LH), which are produced by the anterior pituitary gland in a variety of mammals (Smith et al., 1985). The primary trigger that releases these hypothalamic hormones, which causes the production of testosterone synthesis and spermatogenesis to occur in the testis (and stimulates follicular growth and ovulation in females), in red foxes seems to be related to the photoperiod (Martins et al., 2006). This study also found genes relating to gonadal function and development.

DNAH11 is a candidate gene found that is involved in cilia movement (primarily respiratory cilia) and the inherited disorder Primary ciliary dyskinesia (PCD) (Lucas et al., 2012). However, Schwabe et al., 2008) concluded in their findings that DNAH11 is not involved with severe male infertility, nevertheless Lucas et al., (2012) did find evidence of decreased progressive sperm motility in mice with defective DNAH11. ARMC3 is also a candidate gene found that plays a role in sperm motility. Infertility can arise due to immotile sperm caused by a deficiency in this gene (Liu et al., 2019) such as a frameshift mutation documented in Swedish Red cattle (*Bos taurus*) (Pausch et al., 2016). While DIAPH2 is a

candidate gene found in this study which has been associated with human fertility while also playing a role in ovarian development (Zhang et al., 2016) and the ZFAND3 candidate gene is known to play a role in spermatogenesis (Otake, Endo, & Park, 2011).

In addition to the DIAPH2 gene mentioned above, other genes involved in female reproductive traits has also been found in this study such as PCSK5, ERMP1, and SOHLH2 which have been shown to be involved with ovarian follicle development (Antenos et al., 2011; Cisternino et al., 2013; Kerr et al., 2007; Shin *et al.*, 2017). Furthermore, it has also been shown that the SOHLH2 transcription factor gene is also important in regulating oocyte and spermatogonial differentiation independent of meiosis (Shin et al., 2017). Variants in these reproductive genes may be to various selective pressures.

The red foxes are taken from a range of latitudes in this study and so a range of photoperiods yet there seems to be no clear associations in allele frequencies relating to the daylight hours and these reproductive related candidate genes found in this study. However, it has been shown that the fennec fox (*Vulpes Zerda*) breeding cycle can be inhibited by artificial light representing continuous long days (14L:10D) (Valdespino, Asa, & Bauman, 2005). Additionally, it also seems likely that the crab-eating fox (*Cerdocyon thous*) and the bush dog (*Speothos venaticus*) raised with artificial light has the effect of reducing the environmental constraint and decreased their interval time of reproduction (Valdespino, Asa and Bauman, 2005; Porton, Kleiman and Rodden, 1987). Furthermore, while wild grey wolves typically reproduce once a year (K. Schmidt et al., 2008) the domestic dog (*C. lupus familiaris*) typically reproduces every 6-8 months (Christie and Bell, 1971; Concannon, 2011).

This indicates that the red fox samples taken close to manmade light sources in this study may have had a change in the photoperiod and so the latitudinal rule no longer applies. The artificial urban light may pose a selective pressure or relax an environmental constraint as it

may have in other canids as mentioned above and also in urban relative to rural bird species (Dominoni, Helm, Lehmann, Dowse, & Partecke, 2013).

Another possible cause for selection due to the actions of humans may have occurred from manmade pollution. Pascoe, Zodrow and Greutert, (2014) points to a study finding a decrease in milk production and a decrease in the survival of red fox kits due to fluoride pollution. While Vecoli, Montano, & Andreassi, (2016) suggests there is strong evidence from clinical and experimental studies for air pollution causality of damage to spermiogenesis primarily through genotoxic effects and epigenetic alterations. Additionally, Annamalai & Namasivayam, (2015) points to many studies of EDC pollutants impacting the reproductive system in mammals, fish and amphibians. Of particular interest is the study that found reduced sperm motility (Mocarelli et al., 2008). These reproductive-related genes may have come under large selective pressures due to the relatively recent activities of humans such as pollutants and pesticides some of which have been known to interfere with normal mammalian sex development (Howdeshell *et al.*, 1999; Negri-Cesi *et al.*, 2008). For example, a candidate gene found in this study called DMRT1 is vital for sex differentiation in a wide range of metazoan life but in mammals the Sry gene evolved to take over the role of DMRT1. However, Zhao, Svingen, Ng, & Koopman, (2015) showed in mice that if DMRT1 was expressed at adequate levels at the right place and time then this gene can resume its role and cause testicular differentiation in mammals (masculinization) (Zhao et al., 2015). The reverse is also true, Zarkower et al., (2011) points out that males may become feminized (male-to-female sex reversal) even in adults if Dmrt1 is deleted.

Industrialization is increasing as the human population grows which increases the pollutant levels further still. Air pollution is on the rise and associations with particulates in the air and lung cancer has been suggested (Raaschou-Nielsen et al., 2013) while cancer in wildlife is

increasing in general possibly due to other pollutants such as BPA from plastic (Erren, Zeuß, Steffany, & Meyer-Rochow, 2009).

4.2.8. Cancer

It is interesting to find candidate genes involved with cancer development as around 75% of wild red foxes live less than one year old while 20% live between one to two years old and only 5% live over two years of age and very few live past three years old (Pastoret, 2007). However in captivity Red foxes have been known to live up to 15 years (Pastoret, 2007). The grey fox (*Urocyon cinereoargenteus*) also seems to have a similar life expectancy in the wild with Wood, (1958) documenting less than 1.6% living four years or older in the wild. However, other fox species can live considerably longer in the wild with the closely related (to the red fox) arctic fox having a maximum lifespan of 12 years in the wild (Macdonald & Sillero-Zubiri, 2007). The kit fox (*Vulpes macrotis*) too has been documented to live up to nine years old in the wild although not many live beyond five years of age (Schwartz et al., 2005). Regardless, candidate genes in this study relating to cancer have been found.

Faulty cadherin genes may lose their adhesive properties and lead to tumours forming as cells detach from other surrounding cells and become invasive (Redies, Hertel, & Hübner, 2012). CDH18 has been shown to exhibit a role in suppressing gliomas (Bai et al., 2018) and hence is known as a tumour suppressor. Gliomas are tumours that form within the CNS (Bai et al., 2018). Other tumour suppressor candidate genes found in this study include; CELF2 (Ramalingam, 2012), NKAIN2 (Mao et al., 2016), PTPRK (Bollu, Mazumdar, Savage, & Brown, 2017), PTCH1 (Ozretic et al., 2016), DAB2IP (Valentino et al., 2017), LRRC4 (Peiyao Li, Xu, Li, & Wu, 2014), LRP1B (Beer et al., 2016), CTNNA2 (Fanjul-Fernández et al., 2013), PTPRD (Veeriah et al., 2009), related to various cancers such as lung cancer.

While candidate genes found to play a role as oncogenes include; SMYD3 (Hamamoto et al., 2004), GLT1D1 (Krzeminski et al., 2016), GNA14 (J. Wang et al., 2018), and CTNND2 (Makunin et al., 2014).

Tumour suppressor genes and oncogenes are involved with regulating the cell cycle, if these are abnormally up regulated, down regulated or muted then cancer can form (Levine, Momand, & Finlay, 1991). Due to increasing human activities and their products (pesticides, herbicides, other EDCs) wildlife and human cancer is on the rise (Annamalai and Namasivayam, 2015; Erren *et al.*, 2009). Furthermore wild red foxes have been found to have cancer although this is very rare (Janovsky & Steineck, 1999). It may be that foxes now living close to humans have a selective pressure exerted on genes that correct DNA damage and on cell proliferation genes due to these new manmade genotoxic stresses in populations close to human settlements (Raaschou-Nielsen et al., 2013).

4.3. Conclusion

Using GBS, followed by ascertaining population structure and then combining two outlier detection methods (PCAdapt and BayeScan) I have provided evidence of local adaptation in the red fox's genome across Europe. I found over 100 candidate genes relating to morphology, metabolism, vision, hearing, olfaction, pigmentation, learning, and behaviour in red foxes. Some of these genes have also demonstrated evidence of local adaptation in other Canids and other animals. Although one candidate gene, MUC19, was found with clear allele frequency changes in four Nordic populations which may have been due to a significant parasitic selective pressure. Further investigation is needed to clarify these findings. By using 'Worldclim' (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) in conjunction with the method 'Bayenv,' (Günther & Coop, 2013) a more detailed analysis into allele frequency

changes due to environmental variables can be done. The Worldclim data base provides climatic data over an approximately a one km² global grid system and has been used in studies to help identify climatic roles involved with local adaptions in high altitude enviroments, acoustic signalling, behaviour, and size in grey wolves and other animals (Campbell-Staton, Edwards, & Losos, 2016; Kirschel et al., 2009; Schweizer et al., 2016; Virgós et al., 2011). Bayenv too has been used to successfully identify loci thought to be involved with local adaptations to environmental variables in grey wolves and other species (Blair, Granka, & Feldman, 2014; Cheng et al., 2012; Jones et al., 2012; Keller, Levsen, Olson, & Tiffin, 2012; Schweizer, et al., 2016). Although my study only looked at diversifying selection, there is potential to expand this knowledge by also looking at other types of selection.

To achieve a more holistic understanding of the role selection may have had on a species evolutionary history, I recommend using the outlier loci found in this study for balancing selection by using BayeScan and then combine these results with the Hudson–Kreitman–Aguade (HKA) method (Hudson, Kreitman, & Aguad, 1987). The (HKA) method has been successfully used in previous studies to identify loci and genes presumed to be under balancing selection (especially with pathogen selective pressures) (Hamblin, Thompson, & Di Rienzo, 2002; Kamau & Charlesworth, 2005; Ochola et al., 2010; Tetteh et al., 2009; Weedall & Conway, 2010). The overlapping results from these two methods (BayeScan and HKA) can so be used to detect candidate genes for balancing selection by using the same process as has been done in this study which may provide a clearer picture of a species past.

However, the current study as mentioned above concentrated only on diversifying selection. The results of my findings are likely due to the diverse habitats and broad range of the red fox in which a variety of biotic and abiotic factors may act as strong selective pressures in adaptive trade-offs which differ from region to region. The success in using GBS along with

PCA and two outlier methods (PCAdapt and BayeScan) in identifying signals of selection and providing information on population structure in this study using a species of wild mammal (red fox) may also prove beneficial in other wild natural populations studies.

The relatively easy, convenient, and cheap method used here to identify local adaptations in the red fox may provide a new sustainable way to keep up with the rapidly growing number of studies focusing on local adaptations in wild species of animals (and plants). Therefore, this method can be used for a variety of cost-effective studies in which to gain a deeper understanding of evolution in general and more specifically of the evolutionary history of a specific species. In some instances these methods may also be used even if the particular species genome has not yet been sequenced (Buonaccorsi et al., 2017; Kardos et al., 2015; Russell et al., 2014; Schweizer, *et al.*, 2016) and these methods may be particularly useful for amphibians which have unusually large genomes (Rovelli, Ruiz-gonzález, Vignoli, & Macale, 2019). Important information on genomic adaptation can also be gained in order to maintain wild mammal and other animal populations and their genetic diversity for the future. For example, information could be used to assess the potential ability of a species (via the genome) to adapt to current and a future predicted (via computer models) changing environment due to climate change which can then be used to guide subsequent actions which would be needed to conserve more vulnerable species (Barbosa et al., 2018; Drury et al., 2016; Thomas, Kennington, Evans, Kendrick, & Stat, 2017; Sgrò, Lowe, & Hoffmann, 2011). A further role for these methods to be used in conservation may be applied to identifying genetic characteristics from wild, domesticated, and hybrid populations allowing ‘pure’ wild species to remain ‘wild’ and prevent the genetic extinction of species (Johnson et al., 2015; Randi, 2008; Trouwborst, 2014).

In addition to conserving wild species, the methods used in this study to detect local adaptations may also be extended to be used in domestic or economically important species.

By identifying genetic loci underlying locally adaptive traits a strategy can be formulated to enhance some desired trait (such as adaptations relating to cold environments, high altitudes, disease resistance, arid areas, metabolic adaptions, and much more) by incorporating (breeding) these traits into populations where it may be most beneficial (Barker, 2001; Hansen, 2004; Hoffmann, 2010).

More generally the methods used in this study (by identifying loci that have undergone evolutionary recent positive selection) may also lead to new insights into the evolution of wild species of animals (including humans and their ancestors) and domestic species wild ancestors.

References

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D. R., & Wayne, R. K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences*, *101*(10), 3490–3494. doi: 10.1073/pnas.0306582101

Aguileta, G., Lengelle, J., Marthey, S., Chiapello, H., Rodolphe, F., Gendrault, A., … Giraud, T. (2010). Finding candidate genes under positive selection in Non-model species: Examples of genes involved in host specialization in pathogens. *Molecular Ecology*, *19*(2), 292–306. doi: 10.1111/j.1365-294X.2009.04454.x

Ahmed, Z. M., Riazuddin, S., Aye, S., Ali, R. A., Venselaar, H., Anwar, S., … Friedman, T. B. (2008). Gene structure and mutant alleles of PCDH15: Nonsyndromic deafness DFNB23 and type 1 Usher syndrome. *Human Genetics*, *124*(3), 215–223. doi: 10.1007/s00439-008-0543-3

Airik, R., Schueler, M., Airik, M., Cho, J., Ulanowicz, K. A., Porath, J. D., … Hildebrandt, F. (2016). SDCCAG8 Interacts with RAB Effector Proteins RABEP2 and ERC1 and Is Required for Hedgehog Signaling. *PLOS ONE*, *11*(5), e0156081. doi: 10.1371/journal.pone.0156081

Akey, J. . M. (2009). Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Research*, (206), 711–722. doi: 10.1101/gr.086652.108

Aksoy, I., Utami, K. H., Winata, C. L., Hillmer, A. M., Rouam, S. L., Briault, S., … Cacheux, V. (2017). Personalized genome sequencing coupled with iPSC technology identifies GTDC1 as a gene involved in neurodevelopmental disorders. *Human Molecular Genetics*, *26*(2), 367–382. doi: 10.1093/hmg/ddw393

Alarcon, V. B. (2010). Cell Polarity Regulator PARD6B Is Essential for Trophectoderm Formation in the Preimplantation Mouse Embryo1. *Biology of Reproduction*, *83*(3), 347–358. doi: 10.1095/biolreprod.110.084400

Alberto, F. J., Derory, J., Boury, C., Frigerio, J. M., Zimmermann, N. E., & Kremer, A. (2013). Imprints of natural selection along environmental gradients in phenology-related genes of Quercus petraea. *Genetics*, *195*(2), 495–512. doi: 10.1534/genetics.113.153783

Alkorta-Aranburu, G., Beall, C. M., Witonsky, D. B., Gebremedhin, A., Pritchard, J. K., & Di Rienzo, A. (2012). The Genetic Architecture of Adaptations to High Altitude in Ethiopia. *PLoS Genetics*, *8*(12). doi: 10.1371/journal.pgen.1003110

Amrein, I., & Slomianka, L. (2010). A morphologically distinct granule cell type in the dentate gyrus of the red fox correlates with adult hippocampal neurogenesis. Brain Research, 1328, 12-24. doi: 10.1016/j.brainres.2010.02.075

Anderson, J. T., Willis, J. H., & Mitchell-olds, T. (2012). Evolutionary genetics of plant adaptation. *Trends in Genetics*, *27*(7), 258–266. doi: 10.1016/j.tig.2011.04.001.Evolutionary

Andres, A. M., Hubisz, M. J., Indap, A., Torgerson, D. G., Degenhardt, J. D., Boyko, A. R., … Nielsen, R. (2009). Targets of Balancing Selection in the Human Genome. *Molecular Biology and Evolution*, *26*(12), 2755–2764. doi: 10.1093/molbev/msp190

Andrews, K. R., Good, J. M., Miller, M. R., & Luikart, G. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Publishing Group*, *17*(2), 81–92. doi: 10.1038/nrg.2015.28

Anitei, M., Ifrim, M., Ewart, M., Cowan, A. E., Carson, J. H., Bansal, R., & Pfeiffer, S. E. (2006). A role for Sec8 in oligodendrocyte morphological differentiation. *Journal of Cell Science*, *119*, 807–818. doi: 0.1242/jcs.02785

Anna Tigano, Tone K Reiertsen, James R Walters, V. L. F. (2018). A complex copy number variant underlies differences in both colour plumage and cold adaptation in a dimorphic seabird. *BioRxiv*, 1–14. doi: 10.1101/507384

Annamalai, J., & Namasivayam, V. (2015). Endocrine disrupting chemicals in the atmosphere: Their effects on humans and wildlife. *Environment International*, *76*, 78–97. doi: 10.1016/j.envint.2014.12.006

Annicchiarico, P., Nazzicari, N., Pecetti, L., Romani, M., Ferrari, B., Wei, Y., & Brummer, E. C. (2017). GBS-Based Genomic Selection for Pea Grain Yield under Severe Terminal Drought. *The Plant Genome*, *10*(2), 1–13. doi: 10.3835/plantgenome2016.07.0072

Antenos, M., Lei, L., Xu, M., Malipatil, A., Kiesewetter, S., & Woodruff, T. K. (2011). Role of PCSK5 Expression in Mouse Ovarian Follicle Development: Identification of the Inhibin α- and β-Subunits as Candidate Substrates. *PLoS ONE*, *6*(3), e17348. doi: 10.1371/journal.pone.0017348

Arno, G., Carss, K. J., Hull, S., Zihni, C., Robson, A. G., Fiorentino, A., … Yu, P. (2017). Biallelic Mutation of ARHGEF18 , Involved in the Determination of Epithelial Apicobasal Polarity, Causes Adult-Onset Retinal Degeneration. *The American Journal of Human Genetics*, *100*(2), 334–342. doi: 10.1016/j.ajhg.2016.12.014

Arnold, B., Corbett-Detig, R. B., Hartl, D., & Bomblies, K. (2013). RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, *22*(11), 3179–3190. doi: 10.1111/mec.12276

Arnyasi, M., Wennerström, L., Jansson, E., Andersson, A., Ryman, N., Chiriboga, F., … Kent, M. P. (2017). Complex genetic diversity patterns of cryptic, sympatric brown trout (Salmo trutta) populations in tiny mountain lakes. *Conservation Genetics*, *18*(5), 1213–1227. doi: 10.1007/s10592-017-0972-4

Asa, C. S., & Valdespino, C. (1998). Canid Reproductive Biology: an Integration of Proximate Mechanisms and Ultimate Causes. *American Zoologist*, *38*(1), 251–259. doi: 10.1093/icb/38.1.251

Asthana, S., Schmidt, S., & Sunyaev, S. (2005). A limited role for balancing selection. *Trends in Genetics*, *21*(1), 30–32. doi: 10.1016/j.tig.2004.11.001

Bai, Y. H., Zhan, Y. B., Yu, B., Wang, W. W., Wang, L., Zhou, J. Q., … Liu, X. Z. (2018). A novel tumor-suppressor, CDH18, inhibits glioma cell invasiveness via UQCRC2 and correlates with the prognosis of glioma patients. *Cellular Physiology and Biochemistry*, *48*(4), 1755–1770. doi: 10.1159/000492317

Bajorek, M., Morita, E., Skalicky, J. J., Morham, S. G., Babst, M., & Sundquist, W. I. (2009). Biochemical Analyses of Human IST1 and Its Function in Cytokinesis. *Molecular Biology of the Cell*, *20*(5), 1360–1373.doi: 10.1091/mbc.e08-05-0475

Baker, L. A., Kirkpatrick, B., Rosa, G. J. M., Gianola, D., Valente, B., Sumner, J. P., … Muir, P. (2017). Genome-wide association analysis in dogs implicates 99 loci as risk variants for anterior cruciate ligament rupture. *PLoS ONE*, *12*(4), 1–19. doi: 10.1371/journal.pone.0173810

Baker, P. J., Dowding, C. V., Molony, S. E., White, P. C. L., & Harris, S. (2007). Activity patterns of urban red foxes (*Vulpes vulpes*) reduce the risk of traffic-induced mortality. *Behavioral Ecology*, *18*(4), 716–724. doi: 10.1093/beheco/arm035

Barbosa, S., Mestre, F., White, T. A., Paupério, J., Alves, P. C., & Searle, J. B. (2018). Integrative

approaches to guide conservation decisions: Using genomics to define conservation units and functional corridors. *Molecular Ecology*, *27*(17), 3452–3465. doi: 10.1111/mec.14806

Barker, J. S. F. (2001). Conservation and management of genetic diversity: a domestic animal perspective. *Canadian Journal of Forest Research*, *31*(4), 588–595. doi: 10.1139/cjfr-31-4-588

Barker, S. (2005). The "flickering switch" of late Pleistocene climate change revisited. *Geophysical Research Letters*, *32*(24), 1–4. doi: 10.1029/2005GL024486

Basto, M. P., Santos-Reis, M., Simoes, L., Grilo, C., Cardoso, L., Cortes, H., … Fernandes, C. (2016). Assessing genetic structure in common but ecologically distinct carnivores: The stone marten and red fox. *PLoS ONE*, *11*(1), 1–26. doi: 10.1371/journal.pone.0145165

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, *13*(4), 969–980. doi: 10.1111/j.1365-294X.2004.02125.x

Beedholm-Ebsen, R., van de Wetering, K., Hardlei, T., Nexo, E., Borst, P., & Moestrup, S. K. (2010). Identification of multidrug resistance protein 1 (MRP1/ABCC1) as a molecular gate for cellular export of cobalamin. *Blood*, *115*(8), 1632–1639. doi: 10.1182/blood-2009-07-232587

Beer, A. G., Zenzmaier, C., Schreinlechner, M., Haas, J., Dietrich, M. F., Herz, J., & Marschang, P. (2016). Expression of a recombinant full-length LRP1B receptor in human non-small cell lung cancer cells confirms the postulated growth-suppressing function of this large LDL receptor family member. *Oncotarget*, *7*(42). doi: 10.18632/oncotarget.11897

Bentzen, B. H., Olesen, S. P., Rønn, L. C. B., & Grunnet, M. (2014). BK channel activators and their therapeutic perspectives. *Frontiers in Physiology*, *5*(OCT), 1–12. doi: 10.3389/fphys.2014.00389

Bersaglieri, T., Sabeti, P. C., Patterson, N., Vanderploeg, T., Schaffner, S. F., Drake, J. A., … Schaffner, S. F. (2004). Genetic Signatures of Strong Recent Positive Selection at the Lactase Gene. *The American Journal of Human Genetics*, *74*(6), 1111–1120. doi: 10.1086/421051

Bilkei-Gorzo, A., & Zimmer, A. (2005). Mutagenesis and Knockout Models: NK1 and Substance P. In *Anxiety and Anxiolytic Drugs* (pp. 143–162). doi: 10.1007/3-540-28082-0_5

Billig, G. M., Pál, B., Fidzinski, P., & Jentsch, T. J. (2011). Ca2+-activated Cl- currents are dispensable for olfaction. *Nature Neuroscience*, *14*(6), 763–769. doi: 10.1038/nn.2821

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics*, *25*(22), 3045–3046. doi: 10.1093/bioinformatics/btp536

Bollu, L. R., Mazumdar, A., Savage, M. I., & Brown, P. H. (2017). Molecular pathways: Targeting protein tyrosine phosphatases in cancer. *Clinical Cancer Research*, *23*(9), 2136–2142. doi: 10.1158/1078-0432.CCR-16-0934

Bourhy, H., Kissi, B., Audry, L., Smreczak, M., Sadkowska-Todys, M., Kulonen, K., … Holmes, E. C. (1999a). Ecology and evolution of rabies virus in Europe. *Journal of General Virology*, *80*(10), 2545–2557. doi: 10.1099/0022-1317-80-10-2545

Bourhy, H., Kissi, B., Audry, L., Smreczak, M., Sadkowska-Todys, M., Kulonen, K., … Holmes, E. C. (1999b). Ecology and evolution of rabies virus in Europe. *Journal of General Virology*, *80*(10), 2545–2557. doi: 10.1099/0022-1317-80-10-2545

Braathen, G. J., Høyer, H., Busk, Ø. L., Tveten, K., Skjelbred, C. F., & Russell, M. B. (2016). Variants in the genes DCTN2 , DNAH10 , LRIG3, and MYO1A are associated with intermediate Charcot-Marie-

Tooth disease in a Norwegian family. *Acta Neurologica Scandinavica*, *134*(1), 67–75. doi: 10.1111/ane.12515

Bunn, H., & Poyton, R. (1996). Oxygen sensing and molecular adaptation to hypoxia. Physiological Reviews, 76(3), 839-885. doi: 10.1152/physrev.1996.76.3.839

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. doi: 10.1093/bioinformatics/btm308

Bradley, C. A., & Altizer, S. (2007). Urbanization and the ecology of wildlife diseases. *Trends in Ecology and Evolution*, *22*(2), 95–102. doi: 10.1016/j.tree.2006.11.001

Briggs, M. M., & Schachat, F. (2002). The superfast extraocular myosin (MYH13) is localized to the innervation zone in both the global and orbital layers of rabbit extraocular muscle. *Journal of Experimental Biology*, *205*(20), 3133–3142.

Buonaccorsi, V. P., Malloy, J., Peterson, M., Brubaker, K., & Grant, C. J. (2017). Population Genomic Analysis of Brook Trout in Pennsylvania's Appalachian Region. *Transactions of the American Fisheries Society*, *146*(3), 485–494. doi: 10.1080/00028487.2017.1285351

Cagnacci, F., Lovari, S., & Meriggi, A. (2003). Carrion dependence and food habits of the red fox in an alpine area. *Italian Journal of Zoology*, *70*(1), 31–38. doi: 10.1080/11250000309356493

Campbell-Staton, S. C., Edwards, S. V., & Losos, J. B. (2016). Climate-mediated adaptation after mainland colonization of an ancestrally subtropical island lizard, Anolis carolinensis. *Journal of Evolutionary Biology*, *29*(11), 2168–2180. doi: 10.1111/jeb.12935

Capurro, M., Izumikawa, T., Suarez, P., Shi, W., Cydzik, M., Kaneiwa, T., … Filmus, J. (2017). Glypican-6 promotes the growth of developing long bones by stimulating Hedgehog signaling. *The Journal of Cell Biology*, *216*(9), jcb.201605119. doi: 10.1083/jcb.201605119

Careau, V., Morand-Ferron, J., & Thomas, D. (2007). Basal Metabolic Rate of Canidae from Hot Deserts to Cold Arctic Climates. *Journal of Mammalogy*, *88*(2), 394–400. doi: 10.1644/06-MAMM-A-111R1.1

Carruthers, V. B., & Suzuki, Y. (2007). Effects of Toxoplasma gondii infection on the brain. *Schizophrenia Bulletin*, *33*(3), 745–751. doi: 10.1093/schbul/sbm008

Carson, M., Meredith, A., Shaw, D., Giotis, E., Lloyd, D., & Loeffler, A. (2012). Foxes As a Potential Wildlife Reservoir formecA-Positive Staphylococci. Vector-Borne And Zoonotic Diseases, 12(7), 583-587. doi: 10.1089/vbz.2011.0825

Cavallini, P. (1995). Variation in the body size of the red fox. *Jstor*, *32*(4), 421–427.

Cen, Z., Jiang, Z., Chen, Y., Zheng, X., Xie, F., Yang, X., … Luo, W. (2018). Intronic pentanucleotide TTTCA repeat insertion in the SAMD12 gene causes familial cortical myoclonic tremor with epilepsy type 1. *Brain*, *141*(8), 2280–2288. doi: 10.1093/brain/awy160

Červený, J., Begall, S., Koubek, P., Nováková, P., & Burda, H. (2011). Directional preference may enhance hunting accuracy in foraging foxes. *Biology Letters*, *7*(3), 355–357. doi: 10.1098/rsbl.2010.1145

Charlesworth, B., & Charlesworth, D. (2017). Population genetics from 1966 to 2016. *Heredity*, *118*(1), 2–9. doi: 10.1038/hdy.2016.55

Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided

populations. *Genetical Research*, *70*(2), S0016672397002954. doi: 10.1017/S0016672397002954

Chase, K., Jones, P., Martin, A., Ostrander, E. A., & Lark, K. G. (2009). Genetic mapping of fixed phenotypes: Disease frequency as a breed characteristic. *Journal of Heredity*, *100*(Supplement 1), S37–S41. doi: 10.1093/jhered/esp011

Cheng, C., White, B. J., Kamdem, C., Mockaitis, K., Costantini, C., Hahn, M. W., & Besansky, N. J. (2012). Ecological genomics of anopheles gambiae along a latitudinal cline: A population-resequencing approach. *Genetics*, *190*(4), 1417–1432. doi: 10.1534/genetics.111.137794

Chen, T., Li, Q., Zhang, X., Long, R., Wu, Y., Wu, J., & Fu, X. (2018). TOX expression decreases with progression of colorectal cancers and is associated with CD4 T-cell density and Fusobacterium nucleatum infection. *Human Pathology*, *79*, 93–101. doi: 10.1016/j.humpath.2018.05.008

Choucair, N., Mignon-Ravix, C., Cacciagli, P., Abou Ghoch, J., Fawaz, A., Mégarbané, A., … Chouery, E. (2015). Evidence that homozygous PTPRD gene microdeletion causes trigonocephaly, hearing loss, and intellectual disability. *Molecular Cytogenetics*, *8*(1), 1–8. doi: 10.1186/s13039-015-0149-0

Christie, D. W., & Bell, E. T. (1971). Some observations on the seasonal incidence and frequency of oestrus in breeding bitches in Britain. *Journal of Small Animal Practice*, *12*(3), 159–167. doi: 10.1111/j.1748-5827.1971.tb06213.x

Chu, J. Q., Shi, G., Fan, Y. M., Choi, I. W., Cha, G. H., Zhou, Y., … Quan, J. H. (2016). Production of IL-1β and inflammasome with Up-regulated expressions of NOD-like receptor related genes in toxoplasma gondii-infected THP-1 macrophages. *Korean Journal of Parasitology*, *54*(6), 711–717. doi: 10.3347/kjp.2016.54.6.711

Cisternino, M., Della Mina, E., Losa, L., Madè, A., Rossetti, G., Bassi, L. A., … Ciccone, R. (2013). Idiopathic Central Precocious Puberty Associated with 11 Mb De Novo Distal Deletion of the Chromosome 9 Short Arm. *Case Reports in Genetics*, *2013*, 1–6. doi: 10.1155/2013/978087

Coates, B. S., Sumerford, D. V., Miller, N. J., Kim, K. S., Sappington, T. W., Siegfried, B. D., & Lewis, L. C. (2009). Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity*, *100*(5), 556–564. doi: 10.1093/jhered/esp028

Concannon, P. W. (2011). Reproductive cycles of the domestic bitch. *Animal Reproduction Science*, *124*(3–4), 200–210. doi: 10.1016/j.anireprosci.2010.08.028

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423. doi: 10.1534/genetics.110.114819

Corsolini, S., Focardi, S., Kannan, K., Tanabe, S., & Tatsukawa, R. (1995). Isomer-specific analysis of polychlorinated biphenyls and 2,3,7,8- tetrachlorodibenzo-p-dioxin equivalents (TEQs) in red fox and human adipose tissue from central Italy. *Archives of Environmental Contamination and Toxicology*, *29*, 61–68.

Cunha, S. R., Le Scouarnec, S., Schott, J.-J., & Mohler, P. J. (2008). Exon organization and novel alternative splicing of the human ANK2 gene: Implications for cardiac function and human cardiac disease. *Journal of Molecular and Cellular Cardiology*, *45*(6), 724–734. doi: 10.1016/j.yjmcc.2008.08.005

Cox, R., Muñoz-Garcia, A., Jurkowitz, M., & Williams, J. (2008). β-Glucocerebrosidase Activity in the Stratum Corneum of House Sparrows following Acclimation to High and Low Humidity. Physiological And Biochemical Zoology, 81(1), 97-105. doi: 10.1086/522652

da Silva Santos, S., Clark, T. G., Campino, S., Suarez-Mutis, M. C., Rockett, K. A., Kwiatkowski, D. P., &

Fernandes, O. (2012). Investigation of Host Candidate Malaria-Associated Risk/Protective SNPs in a Brazilian Amazonian Population. *PLoS ONE*, *7*(5), e36692. doi: 10.1371/journal.pone.0036692

Dalongeville, A., Andrello, M., Mouillot, D., Lobreaux, S., Frida, M. F., Belmaker, J., … Manel, S. (2018). Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. *Evolutionary Applications*, *11*(8), 1437–1447. doi: 10.1111/eva.12638

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Publishing Group*, *12*(7), 499–510. doi: 10.1038/nrg3012

De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. *PLoS ONE*, *8*(5). doi: 10.1371/journal.pone.0062137

De Liberato, C., Berrilli, F., Odorizi, L., Scarcella, R., Barni, M., Amoruso, C., … Scaramozzino, P. (2018). Parasites in stray dogs from Italy: Prevalence, risk factors and management concerns. *Acta Parasitologica*, *63*(1), 27–32. doi: 10.1515/ap-2018-0003

de Mooij-van Malsen, J. G., van Lith, H. A., Laarakker, M. C., Brandys, M. K., Oppelaar, H., Collier, D. A., … Kas, M. J. (2013). Cross-species genetics converge to TLL2 for mouse avoidance behavior and human bipolar disorder. *Genes, Brain and Behavior*, *12*(6), 653–657. doi: 10.1111/gbb.12055

Deak, G., Gherman, C. M., Ionică, A. M., Vezendan, A. D., D'Amico, G., Matei, I. A., … Mihalca, A. D. (2017). Angiostrongylus vasorum in Romania: An extensive survey in red foxes, Vulpes vulpes. *Parasites and Vectors*, *10*(1), 1–6. doi: 10.1186/s13071-017-2270-x

Debenham, J. J., Landuyt, H., Troell, K., Tysnes, K., & Robertson, L. J. (2017). Occurrence of Giardia in Swedish Red Foxes (*Vulpes vulpes*). *Journal of Wildlife Diseases*, *53*(3), 649–652. doi: 10.7589/2017-01-002

Deliu, E., Arecco, N., Morandell, J., Dotter, C. P., Contreras, X., Girardot, C., … Novarino, G. (2018). Haploinsufficiency of the intellectual disability gene SETD5 disturbs developmental gene expression and cognition. *Nature Neuroscience*, *21*(12), 1717–1727. doi: 10.1038/s41593-018-0266-2

Dell'Arte, G. L., Laaksonen, T., Norrdahl, K., & Korpimäki, E. (2007). Variation in the diet composition of a generalist predator, the red fox, in relation to season and density of main prey. *Acta Oecologica*, *31*(3), 276–281. doi: 10.1016/j.actao.2006.12.007

Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, *201*(1), 45–56. doi: 10.1111/nph.12441

Desai, A. A., Arana, N., Kafisanwo, O., Letsiou, E., Warda, A., Chen, J., … Machado, R. F. (2013). N-Acetylgalactosaminyltransferase 13 (galnt13) Is A Novel Candidate Gene In Pulmonary Arterial Hypertension And In Pathobiologic Responses To Hypoxia. *American Journal of Respiratory and Critical Care Medicine*, *13*(187), 98454.

Devier, B., Aguileta, G., Hood, M. E., & Giraud, T. (2009). Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics*, *181*(1), 209–223. doi: 10.1534/genetics.108.093708

Díaz-Ruiz, F., Caro, J., Delibes-Mateos, M., Arroyo, B., & Ferreras, P. (2016a). Drivers of red fox (*Vulpes vulpes*) daily activity: prey availability, human disturbance or habitat structure? *Journal of Zoology*, *298*(2), 128–138. doi: 10.1111/jzo.12294

Díaz-Ruiz, F., Caro, J., Delibes-Mateos, M., Arroyo, B., & Ferreras, P. (2016b). Drivers of red fox (*Vulpes vulpes*) daily activity: prey availability, human disturbance or habitat structure? *Journal of*

Zoology, *298*(2), 128–138. doi: 10.1111/jzo.12294

Dip, R., Stieger, C., Deplazes, P., Hegglin, D., Müller, U., Dafflon, O., … Naegeli, H. (2001). Comparison of heavy metal concentrations in tissues of red foxes from adjacent urban, suburban, and rural areas. *Archives of Environmental Contamination and Toxicology*, *40*(4), 551–556. doi: 10.1007/s002440010209

Dip, Ramiro, Hegglin, D., Deplazes, P., Dafflon, O., Koch, H., & Naegeli, H. (2003). Age- and sex-dependent distribution of persistent organochlorine pollutants in urban foxes. *Environmental Health Perspectives*, *111*(13), 1608–1612. doi: 10.1289/ehp.6226

Ditchkoff, S. S., Saalfeld, S. T., & Gibson, C. J. (2006). Animal behavior in urban ecosystems: Modifications due to human-induced stress. *Urban Ecosystems*, *9*(1), 5–12. doi: 10.1007/s11252-006-3262-3

DiTommaso, T., Jones, L. K., Cottle, D. L., Gerdin, A.-K., Vancollie, V. E., Watt, F. M., … Smyth, I. M. (2014). Identification of Genes Important for Cutaneous Function Revealed by a Large Scale Reverse Genetic Screen in the Mouse. *PLoS Genetics*, *10*(10), e1004705. doi: 10.1371/journal.pgen.1004705

Diviani, D., Maric, D., Pérez López, I., Cavin, S., & del Vescovo, C. D. (2013). A-kinase anchoring proteins: Molecular regulators of the cardiac stress response. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1833*(4), 901–908. doi: 10.1016/j.bbamcr.2012.07.014

Dominoni, D. M., Helm, B., Lehmann, M., Dowse, H. B., & Partecke, J. (2013). Clocks for the city: Circadian differences between forest and city songbirds. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1763). doi: 10.1098/rspb.2013.0593

Dong, J., Wyss, A., Yang, J., Price, T. R., Nicolas, A., Nalls, M., … Chen, H. (2017). Genome-Wide Association Analysis of the Sense of Smell in U.S. Older Adults: Identification of Novel Risk Loci in African-Americans and European-Americans. *Molecular Neurobiology*, *54*(10), 8021–8032. doi: 10.1007/s12035-016-0282-8

dos Santos, G., Schroeder, A. J., Goodman, J. L., Strelets, V. B., Crosby, M. A., Thurmond, J., … Gelbart, W. M. (2015). FlyBase: introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Research*, *43*(D1), D690–D697. https://doi.org/10.1093/nar/gku1099

Douglas, S. D., & Leeman, S. E. (2011). Neurokinin-1 receptor: functional significance in the immune system in reference to selected infections and inflammation. *Annals of the New York Academy of Sciences*, *1217*(1), 83–95. doi: 10.1111/j.1749-6632.2010.05826.x

Dubey, J. . (1998). Advances in the life cycle of Toxoplasma gondii. *International Journal for Parasitology*, *28*(7), 1019–1024. doi: 10.1016/S0020-7519(98)00023-X

Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E., & Blum, M. G. B. (2016). Detecting genomic signatures of natural selection with principal component analysis: Application to the 1000 genomes data. *Molecular Biology and Evolution*, *33*(4), 1082–1093. doi: 10.1093/molbev/msv334

Drury, C., Dale, K. E., Panlilio, J. M., Miller, S. V, Lirman, D., Larson, E. A., … Oleksiak, M. F. (2016). Genomic variation among populations of threatened coral: Acropora cervicornis. *BMC Genomics*, *17*(1), 286. doi: 10.1186/s12864-016-2583-8

Dybing, N. A., Fleming, P. A., & Adams, P. J. (2013a). Environmental conditions predict helminth prevalence in red foxes in Western Australia. *International Journal for Parasitology: Parasites and Wildlife*, *2*(1), 165–172. doi: 10.1016/j.ijppaw.2013.04.004

Dybing, N. A., Fleming, P. A., & Adams, P. J. (2013b). Environmental conditions predict helminth

prevalence in red foxes in Western Australia. *International Journal for Parasitology: Parasites and Wildlife*, *2*(1), 165–172. doi: 10.1016/j.ijppaw.2013.04.004

Eads, B. D., Andrews, J., & Colbourne, J. K. (2008). Ecological genomics in Daphnia: Stress responses and environmental sex determination. *Heredity*, *100*(2), 184–190. doi: 10.1038/sj.hdy.6800999

Ebner, K., & Singewald, N. (2006). The role of substance P in stress and anxiety responses. *Amino Acids*, *31*(3), 251–272. doi: 10.1007/s00726-006-0335-9

Eckert, A. J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., Gonzalez-Martinez, S. C., & Neale, D. B. (2010). Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (Pinus taeda L., Pinaceae). *Genetics*, *185*(3), 969–982. doi: 10.1534/genetics.110.115543

Edwards, C. J., Soulsbury, C. D., Statham, M. J., Ho, S. Y. W., Wall, D., Dolf, G., … Bradley, D. G. (2012). Temporal genetic variation of the red fox, *Vulpes vulpes*, across western Europe and the British Isles. *Quaternary Science Reviews*, *57*, 95–104. doi: 10.1016/j.quascirev.2012.10.010

Ehlers, C. L., Gizer, I. R., Bizon, C., Slutske, W., Peng, Q., Schork, N. J., & Wilhelmsen, K. C. (2016). Single nucleotide polymorphisms in the REG-CTNNA2 region of chromosome 2 and NEIL3 associated with impulsivity in a Native American sample. *Genes, Brain and Behavior*, *15*(6), 568–577. doi: 10.1111/gbb.12297

Eichstaedt, C. A., Antão, T., Pagani, L., Cardona, A., Kivisild, T., & Mormina, M. (2014). The Andean adaptive toolkit to counteract high altitude maladaptation: Genome-wide and phenotypic analysis of the Collas. *PLoS ONE*, *9*(3). doi: 10.1371/journal.pone.0093314

Einum, S., & Fleming, I. A. (2000). Highly fecund mothers sacrifice offspring survival to maximize fitness. *Nature*, *405*(6786), 565–567. doi: 10.1038/35014600

Ekblom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, *107*(1), 1–15. doi: 10.1038/hdy.2010.152

Elias, P. M. (2007). The skin barrier as an innate immune element. *Seminars in Immunopathology*, *29*(1), 3–14. doi: 10.1007/s00281-007-0060-9

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE*, *6*(5), e19379. doi: 10.1371/journal.pone.0019379

Erin Chen, Y., Fischbach, M. A., & Belkaid, Y. (2018). Skin microbiota-host interactions. *Nature*, *553*(7689), 427–436. doi: 10.1038/nature25177

Erren, T., Zeuß, D., Steffany, F., & Meyer-Rochow, B. (2009). Increase of wildlife cancer: An echo of plastic pollution? *Nature Reviews Cancer*, *9*(11), 842. doi: 10.1038/nrc2665-c1

Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285–298. doi: 10.1038/hdy.2009.74

Fabrigoule, C., & Maurel, D. (1982). Radio-tracking study of foxes' movements related to their home range. a cognitive map hypothesis. *The Quarterly Journal of Experimental Psychology Section B*, *34*(4), 195–208. doi: 10.1080/14640748208400871

Falguères, C. (2003). ESR dating and the human evolution: Contribution to the chronology of the earliest humans in Europe. *Quaternary Science Reviews*, *22*(10–13), 1345–1351. doi: 10.1016/S0277-3791(03)00047-7

Fanjul-Fernández, M., Quesada, V., Cabanillas, R., Cadiñanos, J., Fontanil, T., Obaya, Á., … López-Otín, C. (2013). Cell-cell adhesion genes CTNNA2 and CTNNA3 are tumour suppressors frequently mutated in laryngeal carcinomas. *Nature Communications*, *4*(5). doi: 10.1038/ncomms3531

Farid, A., Gardner, K., Butler, L., Rupasinghe, P. P., & Myles, S. (2014). Genome-wide Association Mapping of Response to Infection by the Aleutian Mink Disease Virus. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*. IEEE.

Feenstra, J. M., Kanaya, K., Pira, C. U., Hoffman, S. E., Eppey, R. J., & Oberg, K. C. (2012). Detection of genes regulated by Lmx1b during limb dorsalization. *Development, Growth & Differentiation*, *54*(4), 451–462. doi: 10.1111/j.1440-169X.2012.01331.x

Ferreira Pissarra, M., Olalla Saad, S. T., & Lazarini, M. (2018). ARHGAP20 (Rho GTPase activating protein 20). *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, *11*(9), 1–3. doi: 10.4267/2042/68537

Fijarczyk, A., & Babik, W. (2015). Detecting balancing selection in genomes: Limits and prospects. *Molecular Ecology*, *24*(14), 3529–3545. doi: 10.1111/mec.13226

Fischer, M. C., Foll, M., Excoffier, L., & Heckel, G. (2011). Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (Microtus arvalis). *Molecular Ecology*, *20*(7), 1450–1462. doi: 10.1111/j.1365-294X.2011.05015.x

Fischer, M. C., Foll, M., Heckel, G., & Excoffier, L. (2014). Continental-scale footprint of balancing and positive selection in a small rodent (microtus arvalis). *PLoS ONE*, *9*(11). doi: 10.1371/journal.pone.0112332

Fischer, M. C., Rellstab, C., Tedder, A., Zoller, S., Gugerli, F., Shimizu, K. K., … Widmer, A. (2013). Population genomic footprints of selection and associations with climate in natural populations of Arabidopsis halleri from the Alps. *Molecular Ecology*, *22*(22), 5594–5607. doi: 10.1111/mec.12521

Fitzpatrick, M. C., Keller, S. R., & Lotterhos, K. E. (2018). *Comment on "Genomic signals of selection predict climate-driven population declines in a migratory bird" Science*. *7279*(8), 2–4. doi: 10.1126/science.aat7279

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993. doi: 10.1534/genetics.108.092221

Forchhammer, M. C., & Asferg, T. (2000). Invading parasites cause a structural shift in red fox dynamics. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1445), 779–786. doi: 10.1098/rspb.2000.1071

Foster, F., & Collard, M. (2013). A Reassessment of Bergmann's Rule in Modern Humans. *PLoS ONE*, *8*(8). doi: 10.1371/journal.pone.0072269

Freedman, A. H., Thomassen, H. A., Buermann, W., & Smith, T. B. (2010). Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*, *19*(17), 3773–3788. doi: 10.1111/j.1365-294X.2010.04684.x

Freeman, S. D., Keino-Masu, K., Masu, M., & Ladher, R. K. (2015). Expression of the heparan sulfate 6-O-endosulfatases, Sulf1 and Sulf2, in the avian and mammalian inner ear suggests a role for sulfation during inner ear development. *Developmental Dynamics*, *244*(2), 168–180. doi: 10.1002/dvdy.24223

French, C. A., Jin, X., Campbell, T. G., Gerfen, E., Groszer, M., Fisher, S. E., & Costa, R. M. (2012). An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning.

Molecular Psychiatry, *17*(11), 1077–1085. doi: 10.1038/mp.2011.105

Fuerst, P. G., Bruce, F., Tian, M., Wei, W., Elstrott, J., Feller, M. B., … Burgess, R. W. (2009). DSCAM and DSCAML1 Function in Self-Avoidance in Multiple Cell Types in the Developing Mouse Retina. *Neuron*, *64*(4), 484–497. doi: 10.1016/j.neuron.2009.09.027

Gershovich, P. M., Gershovich, Y. G., & Buravkova, L. B. (2013). Molecular genetic features of human mesenchymal stem cells after their osteogenic differentiation under the conditions of microgravity. *Human Physiology*, *39*(5), 540–544. doi: 10.1134/s036211971305006x

Gheusi, G., Ortega-Perez, I., Murray, K., & Lledo, P. (2009). A niche for adult neurogenesis in social behavior. Behavioural Brain Research, 200(2), 315-322. doi: 10.1016/j.bbr.2009.02.006

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE*, *9*(2), e90346. doi: 10.1371/journal.pone.0090346

Glerup, S., Bolcho, U., Mlgaard, S., Bggild, S., Vaegter, C. B., Smith, A. H., … Nykjaer, A. (2016). SorCS2 is required for BDNF-dependent plasticity in the hippocampus. *Molecular Psychiatry*, *21*(12), 1740–1751. doi: 10.1038/mp.2016.108

Gomes, A. C., & Valente, A. (2016). Cranial and body size variation in the Iberian red fox (*Vulpes vulpes silacea*). *Mammalian Biology*, *81*(6), 638–643. doi: 10.1016/j.mambio.2016.08.005

Goszczyński, J. (2002). Home ranges in red fox: territoriality diminishes with increasing area. *Acta Theriologica*, *47*(1), 103–114. doi: 10.1007/BF03192482

Goszczyński, J., Misiorowska, M., & Juszko, S. (2008). Changes in the density and spatial distribution of red fox dens and cub numbers in central Poland following rabies vaccination. *Acta Theriologica*, *53*(2), 121–127. doi: 10.1007/BF03194245

Grossmann, S., Bauer, S., Robinson, P. N., & Vingron, M. (2007). Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics*, *23*(22), 3024–3031. doi: 10.1093/bioinformatics/btm440

Gumy, L. F., Katrukha, E. A., Grigoriev, I., Jaarsma, D., Kapitein, L. C., Akhmanova, A., & Hoogenraad, C. C. (2017). MAP2 Defines a Pre-axonal Filtering Zone to Regulate KIF1- versus KIF5-Dependent Cargo Transport in Sensory Neurons. *Neuron*, *94*(2), 347-362.e7. doi: 10.1016/j.neuron.2017.03.046

Guo, B., Li, Z., & Merilä, J. (2016). Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Molecular Ecology*, *25*(12), 2833–2852. doi: 10.1111/mec.13657

Gupta, N., Santhosh, S. R., Babu, J. P., Parida, M. M., & Rao, P. V. L. (2010). Chemokine profiling of Japanese encephalitis virus-infected mouse neuroblastoma cells by microarray and real-time RT-PCR: Implication in neuropathogenesis. *Virus Research*, *147*(1), 107–112. doi: 10.1016/j.virusres.2009.10.018

Halder, D., Mandal, C., Lee, B., Lee, J., Choi, M., Chai, J., … Chai, Y. (2015). PCDHB14 - and GABRB1 - like nervous system developmental genes are altered during early neuronal differentiation of NCCIT cells treated with ethanol. *Human & Experimental Toxicology*, *34*(10), 1017–1027. doi: 10.1177/0960327114566827

Halpin, M. A., & Bissonette, J. A. (2008). Influence of snow depth on prey availability and habitat use by red fox. *Canadian Journal of Zoology*, *66*(3), 587–592. doi: 10.1139/z88-086

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F. P., Li, M., … Nakamura, Y. (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nature Cell Biology*, *6*(8), 731–740. doi: 10.1038/ncb1151

Hamblin, M. T., Thompson, E. E., & Di Rienzo, A. (2002). Complex Signatures of Natural Selection at the Duffy Blood Group Locus. *The American Journal of Human Genetics*, *70*(2), 369–383. doi: 10.1086/338628

Hansen, P. . (2004). Physiological and cellular adaptations of zebu cattle to thermal stress. *Animal Reproduction Science*, *82–83*, 349–360. doi: 10.1016/j.anireprosci.2004.04.011

Harris, A. S., Trewhella, W. J., & Harris, B. Y. S. (2015). Population an Analysis of Some of the Factors Affecting Dispersal in an Urban Fox (*Vulpes vulpes*) Population. *Journal of Applied Ecology*, *25*(2), 409–422.

Harris, S. J., Parry, R. V., Westwick, J., & Ward, S. G. (2008). Phosphoinositide Lipid Phosphatases: Natural Regulators of Phosphoinositide 3-Kinase Signaling in T Lymphocytes. *Journal of Biological Chemistry*, *283*(5), 2465–2469. doi: 10.1074/jbc.R700044200

Harris, S., & Rayner, J. (1986). Urban Fox (Vulpes vulpes) Population Estimates and Habitat Requirements in Several British Cities. *Journal of Animal Ecology*, *55*(2), 575–591. Available at: http://www.jstor.org/stable/4740 [Accessed: 06 March 2018].

Hartová-Nentvichová, M., Šálek, M., Červený, J., & Koubek, P. (2010). Variation in the diet of the red fox (*Vulpes vulpes*) in mountain habitats: Effects of altitude and season. *Mammalian Biology*, *75*(4), 334–340. doi: 10.1016/j.mambio.2009.09.003

Hashmi, J. A., Albarry, M. A., Almatrafi, A. M., Albalawi, A. M., Mahmood, A., & Basit, S. (2018). Whole exome sequencing identified a novel single base pair insertion mutation in the EYS gene in a six generation family with retinitis pigmentosa. *Congenital Anomalies*, *58*(1), 10–15. doi: 10.1111/cga.12225

Hata, L., Murakami, M., Kuwahara, K., Nakagawa, Y., Kinoshita, H., Usami, S., … Nakao, K. (2011). Zinc-finger protein 90 negatively regulates neuron-restrictive silencer factor-mediated transcriptional repression of fetal cardiac genes. *Journal of Molecular and Cellular Cardiology*, *50*(6), 972–981. doi: 10.1016/j.yjmcc.2011.01.017

Herman, J. S., Jóhannesdóttir, F., Jones, E. P., Mcdevitt, A. D., Michaux, J. R., White, T. A., … Searle, J. B. (2017). Post-glacial colonization of Europe by the wood mouse, Apodemus sylvaticus: Evidence of a northern refugium and dispersal with humans. *Biological Journal of the Linnean Society*, *120*(2), 313–332. doi: 10.1111/bij.12882

Hermey, G. (2009). The Vps10p-domain receptor family. *Cellular and Molecular Life Sciences*, *66*(16), 2677–2689. doi: 10.1007/s00018-009-0043-1

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*(1–2), 87–112. doi: 10.1006/bijl.1999.0332

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. doi: 10.1002/joc.1276

Hoffmann, I. (2010). Climate change and the characterization, breeding and conservation of animal genetic resources. *Animal Genetics*, *41*, 32–46. doi: 10.1111/j.1365-2052.2010.02043.x

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLoS Genetics*, *6*(2), e1000862. doi: 10.1371/journal.pgen.1000862

Holmala, K., & Kauhala, K. (2006). Ecology of wildlife rabies in Europe. *Mammal Review*, *36*(1), 17–36. doi: 10.1111/j.1365-2907.2006.00078.x

Holmes, R. S. (2017). Comparative and evolutionary studies of mammalian arylsulfatase and sterylsulfatase genes and proteins encoded on the X-chromosome. *Computational Biology and Chemistry*, *68*, 71–77. doi: 10.1016/j.compbiolchem.2017.02.009

Hoppe, J. M., Frick, A., Åhs, F., Linnman, C., Appel, L., Jonasson, M., … Furmark, T. (2018a). Association between amygdala neurokinin-1 receptor availability and anxiety-related personality traits. *Translational Psychiatry*, *8*(1), 1–8. doi: 10.1038/s41398-018-0163-1

Hoppe, J. M., Frick, A., Åhs, F., Linnman, C., Appel, L., Jonasson, M., … Furmark, T. (2018b). Association between amygdala neurokinin-1 receptor availability and anxiety-related personality traits. *Translational Psychiatry*, *8*(1), 1–8. doi: 10.1038/s41398-018-0163-1

Horibata, Y., Ando, H., Itoh, M., & Sugimoto, H. (2013). Enzymatic and transcriptional regulation of the cytoplasmic acetyl-CoA hydrolase ACOT12. *Journal of Lipid Research*, *54*(8), 2049–2059. doi: 10.1194/jlr.M030163

Howard, B. A. (2008). The role of NRG3 in mammary development. *Journal of Mammary Gland Biology and Neoplasia*, *13*(2), 195–203. doi: 10.1007/s10911-008-9082-8

Howdeshell, K. L., Hotchkiss, A. K., Thayer, K. A., Vandenbergh, J. G., & vom Saal, F. S. (1999). Exposure to bisphenol A advances puberty. *Nature*, *401*(6755), 763–764. doi: 10.1038/44517

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., … Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498–503. doi: 10.1038/nature12111

Hsu, K.-L., Pugh, H., Tsuboi, K., Adibekian, A., Cravatt, B. F., & Masuda, K. (2012). DAGLβ inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nature Chemical Biology*, *8*(12), 999–1007. doi: 10.1038/nchembio.1105

Hu, F., Robins, A. M., Paushter, D. H., Smolka, M. B., Sullivan, P. M., Zhou, X., & Kim, D. (2016). The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathologica Communications*, *4*(1), 1–16. doi: 10.1186/s40478-016-0324-5

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. doi: 10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. doi: 10.1038/nprot.2008.211

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009c). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. doi: 10.1038/nprot.2008.211

Huang, M., Piao, S., Janssens, I. A., Zhu, Z., Wang, T., Wu, D., … Peñuelas, J. (2017). Velocity of change in vegetation productivity over northern high latitudes. *Nature Ecology and Evolution*, *1*(11), 1649–1654. doi: 10.1038/s41559-017-0328-y

Huang, Y. F., Poland, J. A., Wight, C. P., Jackson, E. W., & Tinker, N. A. (2014). Using Genotyping-By-Sequencing (GBS) for genomic discovery in cultivated oat. *PLoS ONE*, *9*(7), 1–16. doi: 10.1371/journal.pone.0102448

Hudson, R. R., Bailey, K., Skarecky, D., Kwiatowski, J., Ayala, F. J., & Rose, S. (1994). Positive Selection in the Superoxide Dismutase (. *Genetics*, *1340*(136), 1329–1340.

Huijzer, B., & Vandenberghe, J. (1998). Climatic reconstruction of the Weichselian Pleniglacial in northwestern and central Europe. *Journal of Quaternary Science*, *13*(5), 391–417. doi: 10.1002/(SICI)1099-1417(1998090)13:5<391::AID-JQS397>3.0.CO;2-6

Imai, H., Shoji, H., Ogata, M., Kagawa, Y., Owada, Y., Miyakawa, T., … Katsuyama, Y. (2017). Dorsal Forebrain-Specific Deficiency of Reelin-Dab1 Signal Causes Behavioral Abnormalities Related to Psychiatric Disorders. *Cerebral Cortex*, *27*(7), 3485–3501. doi: 10.1093/cercor/bhv334

Iossa, G., Soulsbury, C. D., Baker, P. J., & Harris, S. (2008). Body mass, territory size, and life-history tactics in a socially monogamous canid, the red fox Vulpes vulpes. *Journal of Mammalogy*, *89*(6), 1481–1490. doi: 10.1644/07-mamm-a-405.1

Janovsky, M., & Steineck, T. (1999). Adenocarcinoma of the mammary gland in a red fox from Austria. *J.Wildl.Dis.*, *35*(0090–3558), 392–394. doi: 10.7589/0090-3558-35.2.392

Jędrzejewski, W., & Jędrzejewska, B. (1992). Foraging and Diet of the Red Fox *Vulpes vulpes* in Relation to Variable Food Resources in Foraging and diet of the red fox *Vulpes vulpes* in relation to variable food resources in Bialowieza National Park , Poland. *Ecography*, *15*(2), 212–220.

Jeffery, R. A., Lankester, M. W., McGrath, M. J., & Whitney, H. G. (2004). Angiostrongylus vasorum and Crenosoma vulpis in red foxes (*Vulpes vulpes*) in Newfoundland, Canada. *Canadian Journal of Zoology*, *82*(1), 66–74. doi: 10.1139/z03-211

Jeong, C., Alkorta-Aranburu, G., Basnyat, B., Neupane, M., Witonsky, D. B., Pritchard, J. K., … Di Rienzo, A. (2014). Admixture facilitates genetic adaptations to high altitude in Tibet. *Nature Communications*, *5*(1), 1–7. doi: 10.1038/ncomms4281

Jha, A. R., Zhou, D., Brown, C. D., Kreitman, M., Haddad, G. G., & White, K. P. (2016). Shared Genetic Signals of Hypoxia Adaptation in Drosophila and in High-Altitude Human Populations. *Molecular Biology and Evolution*, *33*(2), 501–517. doi: 10.1093/molbev/msv248

Ji, L. D., Qiu, Y. Q., Xu, J., Irwin, D. M., Tam, S. C., Tang, N. L. S., & Zhang, Y. P. (2012). Genetic adaptation of the hypoxia-inducible factor pathway to oxygen pressure among eurasian human populations. *Molecular Biology and Evolution*, *29*(11), 3359–3370. doi: 10.1093/molbev/mss144

Johnson, J. L., Wittgenstein, H., Mitchell, S. E., Hyma, K. E., Temnykh, S. V., Kharlamova, A. V., … Kukekova, A. V. (2015). Genotyping-by-sequencing (GBS) detects genetic structure and confirms behavioral QTL in tame and aggressive foxes (*Vulpes vulpes*). *PLoS ONE*, *10*(6), 1–22. doi: 10.1371/journal.pone.0127013

Johnson, M. T. J., & Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, *358*(6363). doi: 10.1126/science.aam8327

Jolles, A. E., Ezenwa, V. O., Etienne, R. S., Turner, W. C., & Olff, H. (2008). Interactions between macroparasites and microparasites drive infection patterns in free-ranging African buffalo. *Ecology*, *89*(8), 2239–2250. doi: 10.1890/07-0995.1

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. doi: 10.1093/bioinformatics/btn129

Jones, F. C., Chan, Y. F., Schmutz, J., Grimwood, J., Brady, S. D., Southwick, A. M., … Kingsley, D. M. (2012). A Genome-wide SNP Genotyping Array Reveals Patterns of Global and Repeated Species-Pair Divergence in Sticklebacks. *Current Biology*, *22*(1), 83–90. doi: 10.1016/j.cub.2011.11.045

Joo, Y. Bin, Lim, J., Tsao, B. P., Nath, S. K., Kim, K., & Bae, S.-C. (2018). Genetic variants in systemic lupus erythematosus susceptibility loci, XKR6 and GLT1D1 are associated with childhood-onset SLE in a Korean cohort. *Scientific Reports*, *8*(1), 9962. doi: 10.1038/s41598-018-28128-z

Josephs, E. B., Berg, J. J., Ross-Ibarra, J., & Coop, G. (2019). Detecting Adaptive Differentiation in Structured Populations with Genomic Data and Common Gardens. *Genetics*, *211*(3), 989–1004. doi: 10.1534/genetics.118.301786

Kamau, E., & Charlesworth, D. (2005). Balancing Selection and Low Recombination Affect Diversity near the Self-Incompatibility Loci of the Plant Arabidopsis lyrata. *Current Biology*, *15*(19), 1773–1778. doi: 10.1016/j.cub.2005.08.062

Kang, S., & Kumanogoh, A. (2013). Semaphorins in bone development, homeostasis, and disease. *Seminars in Cell and Developmental Biology*, *24*(3), 163–171. doi: 10.1016/j.semcdb.2012.09.008

Karakas, B., Bachman, K. E., & Park, B. H. (2006). Mutation of the PIK3CA oncogene in human cancers. *British Journal of Cancer*, *94*(4), 455–459. doi: 10.1038/sj.bjc.6602970

Karamon, J., Dabrowska, J., Kochanowski, M., Samorek-Pieróg, M., Sroka, J., Różycki, M., … Cencek, T. (2018). Prevalence of intestinal helminths of red foxes (*Vulpes vulpes*) in central Europe (Poland): A significant zoonotic threat. *Parasites and Vectors*, *11*(1), 1–10. doi: 10.1186/s13071-018-3021-3

Karasugi, T., Semba, K., Hirose, Y., Kelempisioti, A., Nakajima, M., Miyake, A., … Ikegawa, S. (2009). Association of the tag SNPs in the human SKT gene (KIAA1217) with lumbar disc herniation. *Journal of Bone and Mineral Research*, *24*(9), 1537–1543. doi: 10.1359/jbmr.090314

Kardos, M., Luikart, G., Bunch, R., Dewey, S., Edwards, W., McWilliam, S., … Kijas, J. (2015). Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. *Molecular Ecology*, *24*(22), 5616–5632. doi: 10.1111/mec.13415

Karlsson, E. K., Baranowska, I., Wade, C. M., Salmon Hillbertz, N. H. C., Zody, M. C., Anderson, N., … Lindblad-Toh, K. (2007). Efficient mapping of mendelian traits in dogs through genome-wide association. *Nature Genetics*, *39*(11), 1321–1328. doi: 10.1038/ng.2007.10

Katzmarzyk, P. T., & Leonard, W. R. (1998). Climatic influences on human body size and proportions: Ecological adaptations and secular trends. *American Journal of Physical Anthropology*, *106*(4), 483–503. doi: 10.1002/(SICI)1096-8644(199808)106:4<483::AID-AJPA4>3.0.CO;2-K

Kauhala, K., Laukkanen, P., Rége, I. Von, & Rkge, I. Von. (2007). Summer food composition and food niche overlap of the raccoon dog , red fox and badger in Finland. *Ecography*, *21*(5), 457–463.

Kelai, S., Maussion, G., Noble, F., Boni, C., Ramoz, N., Moalic, J.-M., … Simonneau, M. (2008). Nrxn3 upregulation in the globus pallidus of mice developing cocaine addiction. *NeuroReport*, *19*(7), 751–755. doi: 10.1097/WNR.0b013e3282fda231

Keller, S. R., Levsen, N., Olson, M. S., & Tiffin, P. (2012). Local Adaptation in the Flowering-Time Gene Network of Balsam Poplar, Populus balsamifera L. *Molecular Biology and Evolution*, *29*(10), 3143–3152. doi: 10.1093/molbev/mss121

Kerfeld, C. A., & Scott, K. M. (2011). Using BLAST to Teach "E-value-tionary" Concepts. *PLoS Biology*, *9*(2), e1001014. doi: 10.1371/journal.pbio.1001014

Kerr, B., Rawlings, N. D., Colgin, L., Garcia-Rudaz, C., Tapia, V., Dissen, G. A., … Galimi, F. (2007). Fxna, a novel gene differentially expressed in the rat ovary at the time of folliculogenesis, is required for normal ovarian histogenesis. *Development*, *134*(5), 945–957. doi: 10.1242/dev.02795

Ketterson, E. D., Nolan Jr., V., Casto, J. M., Buerkle, C. A., Clotfelter, E., Grindstaff, J. L., … Snajdr, E. (2001). Testosterone, phenotype and fitness: A research program in evolutionary behavioral endocrinology. *Avian Endocrinology*, 19–40.

Khatkar, M. S., Nicholas, F. W., Collins, A. R., Zenger, K. R., Cavanagh, J. AL, Barris, W., … Raadsma, H. W. (2008). Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics*, *9*(1), 187. doi: 10.1186/1471-2164-9-187

Kibar, Z., Bosoi, C. M., Kooistra, M., Salem, S., Finnell, R. H., De Marco, P., … Gros, P. (2009). Novel mutations in VANGL1 in neural tube defects. *Human Mutation*, *30*(7), E706–E715. doi: 10.1002/humu.21026

Kidawa, D., & Kowalczyk, R. (2011). The effects of sex, age, season and habitat on diet of the red fox *Vulpes vulpes* in northeastern Poland. *Acta Theriologica*, *56*(3), 209–218. doi: 10.1007/s13364-011-0031-3

Kim, J. Y., Choi, S. Y., Moon, Y., Kim, H. J., Chin, J. H., Kim, H., & Sun, W. (2012). Different expression patterns of Phactr family members in normal and injured mouse brain. *Neuroscience*, *221*, 37–46. doi: 10.1016/j.neuroscience.2012.06.059

Kimbrell, D. A., & Beutler, B. (2001). The evolution and genetics of innate immunity. *Nature Reviews Genetics*, *2*(4), 256–267. doi: 10.1038/35066006

Kirschel, A. N. G., Blumstein, D. T., Cohen, R. E., Buermann, W., Smith, T. B., & Slabbekoorn, H. (2009). Birdsong tuned to the environment: green hylia song varies with elevation, tree cover, and noise. *Behavioral Ecology*, *20*(5), 1089–1095. doi: 10.1093/beheco/arp101

Klinger, M., Wang, W., Kuhns, S., Bärenz, F., Dräger-Meurer, S., Pereira, G., & Gruss, O. J. (2014). The novel centriolar satellite protein SSX2IP targets Cep290 to the ciliary transition zone. *Molecular Biology of the Cell*, *25*(4), 495–507. doi: 10.1091/mbc.e13-09-0526

Kronqvist, P., Kawaguchi, N., Albrechtsen, R., Xu, X., Schrøder, H. D., Moghadaszadeh, B., … Wewer, U. M. (2002). ADAM12 Alleviates the Skeletal Muscle Pathology in. *American Journal of Pathology, 161*(5), 1535–1540. doi: 10.1016/S0002-9440(10)64431-8

Krzeminski, P., Corchete, L. A., García, J. L., López-Corral, L., Fermiñán, E., García, E. M., … Gutiérrez, N. C. (2016). Integrative analysis of DNA copy number, DNA methylation and gene expression in multiple myeloma reveals alterations related to relapse. *Oncotarget*, *7*(49), 80664–80679. doi: 10.18632/oncotarget.13025

Kukekova, A. V., Johnson, J. L., Xiang, X., Feng, S., Liu, S., Rando, H. M., … Zhang, G. (2018). Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nature Ecology and Evolution*, *2*(9), 1479–1491. doi: 10.1038/s41559-018-0611-6

Kurki Sami, Nikula Ari, Helle Pekka, L. H. (1998). Abundances of Red Fox and Pine Marten in Relation to the Composition of Boreal Forest Landscapes. *The Journal of Animal Ecology*, *67*(6), 874–886.

Kutschera, V. E., Lecomte, N., Janke, A., Selva, N., Sokolov, A. A., Haun, T., … Hailer, F. (2013). A range-wide synthesis and timeline for phylogeographic events in the red fox (*Vulpes vulpes*). *BMC*

Evolutionary Biology, *13*(1). doi: 10.1186/1471-2148-13-114

Lai, T.-H., Wu, Y.-Y., Wang, Y.-Y., Chen, M.-F., Wang, P., Chen, T.-M., … Lin, Y.-H. (2016). SEPT12–NDC1 Complexes Are Required for Mammalian Spermiogenesis. *International Journal of Molecular Sciences*, *17*(11), 1911. doi: 10.3390/ijms17111911

Lamb, T. D., Patel, H. R., Chuah, A., & Hunt, D. M. (2018). Evolution of the shut-off steps of vertebrate phototransduction. *Open Biology*, *8*(1), 170232. doi: 10.1098/rsob.170232

Laurenti, E., Frelin, C., Xie, S., Ferrari, R., Dunant, C. F., Zandi, S., … Dick, J. E. (2015). CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell*, *16*(3), 302–313. doi: 10.1016/j.stem.2015.01.017

Laurimaa, L., Süld, K., Moks, E., Valdmann, H., Umhang, G., Knapp, J., & Saarma, U. (2015). First report of the zoonotic tapeworm Echinococcus multilocularis in raccoon dogs in Estonia, and comparisons with other countries in Europe. *Veterinary Parasitology*, *212*(3–4), 200–205. doi: 10.1016/j.vetpar.2015.06.004

Lempp, C., Jungwirth, N., Grilo, M. L., Reckendorf, A., Ulrich, A., van Neer, A., … Siebert, U. (2017). Pathological findings in the red fox (*Vulpes vulpes*), stone marten (Martes foina) and raccoon dog (Nyctereutes procyonoides), with special emphasis on infectious and zoonotic agents in Northern Germany. *PLOS ONE*, *12*(4), 1–20. doi: 10.1371/journal.pone.0175469

Letková, V., Lazar, P., Urlík, J., Goldová, M., Kocišová, A., & Košuthová, L. (2006). The red fox (*Vulpes vulpes L*.) as a source of zoonoses. *Vet Archiv*, *76*, S73–S81.

Levine, A. J., Momand, J., & Finlay, C. A. (1991). The p53 tumour suppressor gene. *Nature*, *351*(6326), 453–456. doi: 10.1038/351453a0

Li, C., Waldbieser, G., Bosworth, B., Beck, B. H., Thongda, W., & Peatman, E. (2014). SNP discovery in wild and domesticated populations of blue catfish, Ictalurus furcatus, using genotyping-by-sequencing and subsequent SNP validation. *Molecular Ecology Resources*, *14*(6), 1261–1270. doi: 10.1111/1755-0998.12272

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(1), 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, M., & Zhao, C. (2009). Study on Tibetan Chicken embryonic adaptability to chronic hypoxia by revealing differential gene expression in heart tissue. *Science in China Series C: Life Sciences*, *52*(3), 284–295. doi: 10.1007/s11427-009-0005-8

Li, Peiyao, Xu, G., Li, G., & Wu, M. (2014). Function and mechanism of tumor suppressor gene LRRC4/NGL-2. *Molecular Cancer*, *13*(1), 266. doi: 10.1186/1476-4598-13-266

Li, Peng, Huang, J., Tian, H. jun, Huang, Q. yuan, Jiang, C. hua, & Gao, Y. qi. (2011). Regulation of bone marrow hematopoietic stem cell is involved in high-altitude erythrocytosis. *Experimental Hematology*, *39*(1), 37–46. doi: 10.1016/j.exphem.2010.10.006

Lightfoot, J. T. (2013). Why Control Activity? Evolutionary Selection Pressures Affecting the Development of Physical Activity Genetic and Biological Regulation. *BioMed Research International*, *2013*, 1–10. doi: 10.1155/2013/821678

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., … Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, *438*(7069), 803–819. doi: 10.1038/nature04338

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. doi: 10.1093/bioinformatics/btr642

Lisowski, P., Stankiewicz, A. M., Goscik, J., Wieczorek, M., Zwierzchowski, L., & Swiergiel, A. H. (2012). Selection for stress-induced analgesia affects the mouse hippocampal transcriptome. *Journal of Molecular Neuroscience*, *47*(1), 101–112. doi: 10.1007/s12031-011-9692-2

Liu, C. X., Musco, S., Lisitsina, N. M., Yaklichkin, S. Y., & Lisitsyn, N. A. (2000). Genomic organization of a new candidate tumor suppressor gene, LRP1B. *Genomics*, *69*(2), 271–274. doi: 10.1006/geno.2000.6331

Liu, T., Kong, W. X., Tang, X. Y., Xu, M., Wang, Q. H., Zhang, B., … Chen, H. (2018). The transcription factor Zfp90 regulates the self-renewal and differentiation of hematopoietic stem cells article. *Cell Death and Disease*, *9*(6). doi: 10.1038/s41419-018-0721-8

Lockhart, P. J., O'Farrell, C. A., & Farrer, M. J. (2004). It's a double knock-out! The quaking mouse is a spontaneous deletion of parkin and parkin co-regulated gene (PACRG). *Movement Disorders*, *19*(1), 101–104. doi: 10.1002/mds.20000

Long, Q., Rabanal, F. A., Meng, D., Huber, C. D., Farlow, A., Platzer, A., … Nordborg, M. (2013). Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. *Nature Genetics*, *45*(8), 884–890. doi: 10.1038/ng.2678

Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K. A., Germann, P., … Zeller, R. (2014). Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature*, *511*(7507), 46–51. doi: 10.1038/nature13289

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, *23*(9), 2178–2192. doi: 10.1111/mec.12725

Lowell, B. B., & Spiegelman, B. M. (2000). Towards a molecular understanding of adaptive thermogenesis. *Nature*, *404*(6778), 652–660. doi: 10.1038/35007527

Lu, X., Liu, S., Wang, H.-H., Yu, F., Liu, J.-J., Zhao, Y., & Zhao, S.-L. (2019). A biological study of supernumerary teeth derived dental pulp stem cells based on RNA-seq analysis. *International Endodontic Journal*, *52*(6), 819–828. doi: 10.1111/iej.13060

Lucas, J. S., Adam, E. C., Goggin, P. M., Jackson, C. L., Powles-Glover, N., Patel, S. H., … Lackie, P. M. (2012). Static respiratory cilia associated with mutations in Dnahc11/DNAH11: A mouse model of PCD. *Human Mutation*, *33*(3), 495–503. doi: 10.1002/humu.22001

Luniak, M. (2004). Synurbization - adaptation of animal wildlife to urban development. *Proceedings Of The 4th International Symposium On Urban Wildlife Conservation*, 50–55. doi: 10.1099/mic.0.069724-0

Luo, R., Reed, C. E., Sload, J. A., Wordeman, L., Randazzo, P. A., & Chen, P.-W. (2019). Arf GAPs and molecular motors. *Small GTPases*, *10*(3), 196–209. doi: 10.1080/21541248.2017.1308850

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77. doi:

10.1111/1755-0998.12592

Macdonald, D., Brown, L., Yerli, S., & Canbolat, A. (1994). Behavior of Red Foxes, Vulpes vulpes, Caching Eggs of Loggerhead Turtles, Caretta caretta. Journal Of Mammalogy, 75(4), 985-988. doi: 10.2307/1382480

Macdonald, D. (2010). Food Caching by Red Foxes and Some Other Carnivores. Zeitschrift Für Tierpsychologie, 42(2), 170-185. doi: 10.1111/j.1439-0310.1976.tb00963.x

Macdonald, D. (1979). 'Helpers'in fox society. Nature, 282(5734), 69-71. doi: 10.1038/282069a0

Macdonald, D. W. (1980). The Red Fox, *Vulpes vulpes*, as a Predator upon Earthworms, Lunzbricus terrestris. *Ethology*, *52*(2), 171–200. doi: 10.1016/b978-0-7020-3008-6.50029-6

Macdonald, D., & Sillero-Zubiri, C. (2007). *The biology and conservation of wild canids* (1st ed., pp. 166-167). Oxford: Oxford Univ. Press.

Mackenstedt, U., Jenkins, D., & Romig, T. (2015). The role of wildlife in the transmission of parasitic zoonoses in peri-urban and urban areas. *International Journal for Parasitology: Parasites and Wildlife*, *4*(1), 71–79. doi: 10.1016/j.ijppaw.2015.01.006

Mackey, M. C. (2001). Cell kinetic status of haematopoietic stem cells. *Cell Proliferation*, *34*(2), 71–83. doi: 10.1046/j.1365-2184.2001.00195.x

Makunin, A. I., Dementyeva, P. V, Graphodatsky, A. S., Volobouev, V. T., Kukekova, A. V, & Trifonov, V. A. (2014). Genes on B chromosomes of vertebrates. *Molecular Cytogenetics*, *7*(1), 99. doi: 10.1186/s13039-014-0099-y

Malczewski, A., Gawor, J., & Malczewska, M. (2008). Infection of red foxes (Vulpes vulpes) with Echinococcus multilocularis during the years 2001-2004 in Poland. *Parasitology Research*, *103*(3), 501–505. doi: 10.1007/s00436-008-0990-8

Malkemper, E. P., Topinka, V., & Burda, H. (2015). A behavioral audiogram of the red fox (*Vulpes vulpes*). *Hearing Research*, *320*, 30–37. doi: 10.1016/j.heares.2014.12.001

Malmasi, A., Khosravi, A. R., Selk Ghaffari, M., & Shojaee Tabrizi, A. (2009). Microsporum canis infection in a red fox (*Vulpes vulpes*). *Iranian Journal of Veterinary Research*, *10*(2), 189–191.

Mao, X., Luo, F., Boyd, L. K., Zhou, B., Zhang, Y., Stankiewicz, E., … Lu, Y.-J. (2016). *NKAIN2* functions as a novel tumor suppressor in prostate cancer. *Oncotarget*, *7*(39), 63793–63803. doi: 10.18632/oncotarget.11690

Martins, M. I. M., Souza, F. F. de, Oba, E., & Lopes, M. D. (2006). The effect of season on serum testosterone concentrations in dogs. *Theriogenology*, *66*(6–7), 1603–1605. doi: 10.1016/j.theriogenology.2006.02.028

Matin, A., & Nadeau, J. H. (2001). Sensitized polygenic trait analysis. *Trends in Genetics*, *17*(12), 727–731.

McCallum, H. (2012). Disease and the dynamics of extinction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1604), 2828–2839. doi: 10.1098/rstb.2012.0224

McDaniel, A. H., Li, X., Tordoff, M. G., Bachmanov, A. A., & Reed, D. R. (2006). A locus on mouse Chromosome 9 (Adip5) affects the relative weight of the gonadal but not retroperitoneal adipose depot. *Mammalian Genome*, *17*(11), 1078–1092. doi: 10.1007/s00335-006-0055-1

McDevitt, A. D., Zub, K., Kawałko, A., Oliver, M. K., Herman, J. S., & Wójcik, J. M. (2012). Climate and refugial origin influence the mitochondrial lineage distribution of weasels (Mustela nivalis) in a

phylogeographic suture zone. *Biological Journal of the Linnean Society*, *106*(1), 57–69. doi: 10.1111/j.1095-8312.2012.01840.x

McDonald, D. W. (1979). Some observation and field experiments on the urine marking behaviour of the red fox, *Vulpes vulpes*. *L. Z. Tierpsychol*, *51*, 1–22.

McKenzie, C. W., Craige, B., Kroeger, T. V., Finn, R., Wyatt, T. A., Sisson, J. H., … Lee, L. (2015). CFAP54 is required for proper ciliary motility and assembly of the central pair apparatus in mice. *Molecular Biology of the Cell*, *26*(18), 3140–3149. doi: 10.1091/mbc.e15-02-0121

McLachlan, G. J. (1999). Mahalanobis distance. *Resonance*, *4*(6), 20–26. doi: 10.1007/bf02834632

McNamee, C. J., Reed, J. E., Howard, M. R., Lodge, A. P., & Moss, D. J. (2002). Promotion of neuronal cell adhesion by members of the IgLON family occurs in the absence of either support or modification of neurite outgrowth. *Journal of Neurochemistry*, *80*(6), 941–948. doi: 10.1046/j.0022-3042.2002.00798.x

McVean, G. (2009). A Genealogical Interpretation of Principal Components Analysis. *PLoS Genetics*, *5*(10), e1000686. doi: 10.1371/journal.pgen.1000686

Medina, A. I., Martí, D. A., & Bidau, C. J. (2007). Subterranean rodents of the genus Ctenomys (Caviomorpha, Ctenomyidae) follow the converse to Bergmann's rule. *Journal of Biogeography*, *34*(8), 1439–1454. doi: 10.1111/j.1365-2699.2007.01708.x

Mercati, O., Huguet, G., Danckaert, A., Maruani, A., Bellinzoni, M., Rolland, T., & Gouder, L. (2017). CNTN6 mutations are risk factors for abnormal auditory sensory perception in autism spectrum disorders. *Molecular Psychiatry*, *22*, 625–633. doi: 10.1038/mp.2016.61

Mocarelli, P., Crespi, C., Turner, W. E., Milani, S., Sarto, C., Carreri, V., … Colombo, L. (2008). Dioxin Exposure, from Infancy through Puberty, Produces Endocrine Disruption and Affects Human Semen Quality. *Environmental Health Perspectives*, *116*(1), 70–77. doi: 10.1289/ehp.10399

Molina, L., Fasquelle, L., Nouvian, R., Salvetat, N., Scott, H. S., Guipponi, M., … Delprat, B. (2013). Tmprss3 loss of function impairs cochlear inner hair cell kcnma1 channel membrane expression. *Human Molecular Genetics*, *22*(7), 1289–1299. doi: 10.1093/hmg/dds532

Møller, A. P., & Szép, T. (2011). The role of parasites in ecology and evolution of migration and migratory connectivity. *Journal of Ornithology*, *152*, S141–S150. doi: 10.1007/s10336-010-0621-x

Montes, M., Cloutier, A., Sanchez-Hernandez, N., Michelle, L., Lemieux, B., Blanchette, M., … Sune, C. (2011). TCERG1 Regulates Alternative Splicing of the Bcl-x Gene by Modulating the Rate of RNA Polymerase II Transcription. *Molecular and Cellular Biology*, *32*(4), 751–762. doi: 10.1128/mcb.06255-11

Munekata, K., & Sakamoto, K. (2009). Forkhead transcription factor Foxo1 is essential for adipocyte differentiation. *In Vitro Cellular and Developmental Biology - Animal*, *45*(10), 642–651. doi: 10.1007/s11626-009-9230-5

Muray, D. L., & Larivière, S. (2002). The relationship between foot size of wild canids and regional snow conditions: Evidence for selection against a high footload? *Journal of Zoology*, *256*(3), 289–299. doi: 10.1017/S095283690200033X

Murcia, C. L., & Woychik, R. P. (2001). Expression of Pcdh15 in the inner ear, nervous system and various epithelia of the developing embryo. *Mechanisms of Development*, *105*(1–2), 163–166. doi: 10.1016/S0925-4773(01)00388-4

Nadachowski, A., & Sommer, R. S. (2006). Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Review*, *36*(4), 251–265. doi: 10.1111/j.1365-2907.2006.00093.x

Nakagawa, N., Yagi, H., Kato, K., Takematsu, H., & Oka, S. (2015). Ectopic clustering of Cajal–Retzius and subplate cells is an initial pathological feature in Pomgnt2-knockout mice, a model of dystroglycanopathy. *Scientific Reports*, *5*(1), 11163. doi: 10.1038/srep11163

Nakayama, K., & Iwamoto, S. (2017). An adaptive variant of TRIB2, rs1057001, is associated with higher expression levels of thermogenic genes in human subcutaneous and visceral adipose tissues. *Journal of Physiological Anthropology*, *36*(1), 16. doi: 10.1186/s40101-017-0132-z

Namba, A., Abe, S., Shinkawa, H., Kimberling, W. J., & Usami, S. (2001). Genetic features of hearing loss associated with ear anomalies: PDS and EYA1 mutation analysis. *Journal of Human Genetics*, *46*(9), 518–521. doi: 10.1007/s100380170033

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*(1), 184–194. doi: 10.1111/j.1755-0998.2011.02987.x

Negri-Cesi, P., Colciago, A., Pravettoni, A., Casati, L., Conti, L., & Celotti, F. (2008). Sexual differentiation of the rodent hypothalamus: Hormonal and environmental influences. *Journal of Steroid Biochemistry and Molecular Biology*, *109*(3–5), 294–299. doi: 10.1016/j.jsbmb.2008.03.003

Neureither, F., Ziegler, K., Pitzer, C., Frings, S., & Möhrlen, F. (2017). Impaired Motor Coordination and Learning in Mice Lacking Anoctamin 2 Calcium-Gated Chloride Channels. *Cerebellum*, *16*(5–6), 929–937. doi: 10.1007/s12311-017-0867-4

Nielsen, R. (2005). Molecular Signatures of Natural Selection. *Annual Review of Genetics*, *39*(1), 197–218. doi: 10.1146/annurev.genet.39.073003.112420

Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., & Clark, A. G. (2007). Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, *8*(11), 857–868. doi: 10.1038/nrg2187

Nishimura, T., Katsumura, T., Motoi, M., Oota, H., & Watanuki, S. (2017). Experimental evidence reveals the UCP1 genotype changes the oxygen consumption attributed to non-shivering thermogenesis in humans. *Scientific Reports*, *7*(1), 1–7. doi: 10.1038/s41598-017-05766-3

Nuytens, K., Gantois, I., Stijnen, P., Iscru, E., Laeremans, A., Serneels, L., … D'Hooge, R. (2013). Haploinsufficiency of the autism candidate gene Neurobeachin induces autism-like behaviors and affects cellular and molecular processes of synaptic plasticity in mice. *Neurobiology of Disease*, *51*, 144–151. doi: 10.1016/j.nbd.2012.11.004

O'Rourke, T., & Boeckx, C. (2018). Converging roles of glutamate receptors in domestication and prosociality. *BioRxiv*, 439869. doi: 10.1101/439869

Oakeshott, J. G., Wilson, S. R., & Knibb, W. R. (1998). Selection affecting enzyme polymorphisms in enclosed Drosophila populations maintained in a natural environment. *Proceedings of the National Academy of Sciences*, *85*(1), 293–297. doi: 10.1073/pnas.85.1.293

Ochola, L. I., Tetteh, K. K. A., Stewart, L. B., Riitho, V., Marsh, K., & Conway, D. J. (2010). Allele frequency-based and polymorphism-versus-divergence indices of balancing selection in a new filtered set of polymorphic genes in Plasmodium falciparum. *Molecular Biology and Evolution*, *27*(10), 2344–2351. doi: 10.1093/molbev/msq119

Oda, S., Nozawa, T., Nozawa-Minowa, A., Tanaka, M., Aikawa, C., Harada, H., & Nakagawa, I. (2016). Golgi-resident GTPase Rab30 promotes the biogenesis of pathogen-containing autophagosomes. *PLoS ONE*, *11*(1), 1–15. doi: 10.1371/journal.pone.0147061

Oh, S.-H., Kim, J.-W., Kim, Y., Lee, M. N., Kook, M.-S., Choi, E. Y., … Koh, J.-T. (2017). The extracellular matrix protein Edil3 stimulates osteoblast differentiation through the integrin α5β1/ERK/Runx2 pathway. *PLOS ONE*, *12*(11), e0188749. doi: 10.1371/journal.pone.0188749

Ohye, T., Inagaki, H., Ozaki, M., Ikeda, T., & Kurahashi, H. (2014). Signature of backward replication slippage at the copy number variation junction. *Journal of Human Genetics*, *59*(5), 247–250. doi: 10.1038/jhg.2014.20

Oleksyk, T. K., Smith, M. W., & O'Brien, S. J. (2010). Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 185–205. doi: 10.1098/rstb.2009.0219

Ong, D. S. T., Wang, Y.-J., Tan, Y. L., Yates, J. R., Mu, T.-W., & Kelly, J. W. (2013). FKBP10 Depletion Enhances Glucocerebrosidase Proteostasis in Gaucher Disease Fibroblasts. *Chemistry & Biology*, *20*(3), 403–415. doi: 10.1016/j.chembiol.2012.11.014

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, *10*(8), 531–539. doi: 10.1038/nrg2603

Otake, S., Endo, D., & Park, M. K. (2011). Molecular characterization of two isoforms of ZFAND3 cDNA from the Japanese quail and the leopard gecko, and different expression patterns between testis and ovary. *Gene*, *488*(1–2), 23–34. doi: 10.1016/j.gene.2011.08.021

Ozretic, P., Bisio, A., Musani, V., Trnski, D., Sabol, M., Inga, A., & Levanat, S. (2016). Expression of PTCH1b tumor suppressor gene is controlled by different 5'-untranslated region cis-regulatory elements. *European Journal of Cancer*, *61*, S164. doi: 10.1016/S0959-8049(16)61578-2

Pagh, S., Hansen, M. S., Jensen, B., Pertoldi, C., & Chriél, M. (2018). Variability in body mass and sexual dimorphism in Danish red foxes (*Vulpes vulpes*) in relation to population density. *Zoology and Ecology*, *28*(1), 1–9. doi: 10.1080/21658005.2017.1409997

Paradis, E. (2018). Reading Genetic Data Files Into R with adegenet and pegas. *Software User Manual*, 1–7.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. doi: 10.1093/bioinformatics/bty633

Pascoe, G. A., Zodrow, J., & Greutert, E. (2014). Evaluating Risks to Terrestrial Wildlife from Environmental Fluoride. *Human and Ecological Risk Assessment*, *20*(4), 941–961. doi: 10.1080/10807039.2012.750162

Pastoret, P.-P. (2007). Challenges and Issues of Early Life Vaccination in Animals and Humans. *Journal of Comparative Pathology*, *137*, S2–S3. doi: 10.1016/j.jcpa.2007.04.003

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, *2*(12), 2074–2093. doi: 10.1371/journal.pgen.0020190

Pausch, H., Venhoranta, H., Wurmser, C., Hakala, K., Iso-Touru, T., Sironen, A., … Andersson, M. (2016). A frameshift mutation in ARMC3 is associated with a tail stump sperm defect in Swedish Red (Bos taurus) cattle. *BMC Genetics*, *17*(1), 1–9. doi: 10.1186/s12863-016-0356-7

Perea, C., Fernando, J., Hoz, D. La, Cruz, D. F., Lobaton, J. D., Izquierdo, P., … Duitama, J. (2016). *Bioinformatic analysis of genotype by sequencing (GBS) data with NGSEP*. *17*(Suppl 5). doi: 10.1186/s12864-016-2827-7

Perrone, M. D., Gerard, M., Reggiani, C., Niel-Bütschi, F., Casimir, G., Deconinck, N., … Belligni, E. F. (2017). Novel promoters and coding first exons in DLG2 linked to developmental disorders and intellectual disability. *Genome Medicine*, *9*(1), 1–20. doi: 10.1186/s13073-017-0452-y

Peterson, L. W., Philip, N. H., Delaney, A., Dolfi, M. A. W., Asklof, K., Gray, F., … Brodsky, I. E. (2017).

RIPK1-dependent apoptosis bypasses pathogen blockade of innate signaling to promote immune defense. *Journal of Experimental Medicine*, *214*(11), 3171–3182. doi: 10.1084/jem.20170347

Petit, F., Sears, K. E., & Ahituv, N. (2017). Limb development: a paradigm of gene regulation. *Nature Reviews Genetics*, *18*(4), 245–258. doi: 10.1038/nrg.2016.167

Pilot, M., Greco, C., Vonholdt, B. M., Jędrzejewska, B., Randi, E., Jędrzejewski, W., … Wayne, R. K. (2014). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity*, *112*(4), 428–442. doi: 10.1038/hdy.2013.122

Pimsler, M. L., Jackson, J. M., & Lozier, J. D. (2017). Population genomics reveals a candidate gene involved in bumble bee pigmentation. *Ecology and Evolution*, *7*(10), 3406–3413. doi: 10.1002/ece3.2935

Plumer, L., Davison, J., Saarma, U., & Cameron, E. Z. (2014). Rapid urbanization of red foxes in estonia: distribution, behaviour, attacks on domestic animals, and health-risks related to zoonotic diseases. *PLoS ONE*, *9*(12), 1–15. doi: 10.1371/journal.pone.0115124

Plump, A. S., Erskine, L., Sabatier, C., Brose, K., Epstein, C. J., Goodman, C. S., … Tessier-Lavigne, M. (2002). Slit1 and Slit2 Cooperate to Prevent Premature Midline Crossing of Retinal Axons in the Mouse Visual System. *Neuron*, *33*(2), 219–232. doi: 10.1016/S0896-6273(01)00586-4

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. *The Plant Genome Journal*, *5*(3), 92. doi: 10.3835/plantgenome2012.05.0005

Pongrácz, P., Ujvári, V., Faragó, T., Miklósi, Á., & Péter, A. (2017). Do you see what I see? The difference between dog and human visual perception may affect the outcome of experiments. *Behavioural Processes*, *140*(November 2016), 53–60. doi: 10.1016/j.beproc.2017.04.002

Pootakham, W., Jomchai, N., Ruang-areerate, P., Shearman, J. R., Sonthirod, C., Sangsrakru, D., … Tangphatsornruang, S. (2015). Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics*, *105*(5–6), 288–295. doi: 10.1016/j.ygeno.2015.02.002

Porton, I. J., Kleiman, D. G., & Rodden, M. (1987). Aseasonality of Bush Dog Reproduction and the Influence of Social Factors on the Estrous Cycle. *Journal of Mammalogy*, *68*(4), 867–871. doi: 10.2307/1381569

Poulsen, N. A., Hemmer-Hansen, J., Loeschcke, V., Carvalho, G. R., & Nielsen, E. E. (2011). Microgeographical population structure and adaptation in Atlantic cod Gadus morhua: Spatio-temporal insights from gene-associated DNA markers. *Marine Ecology Progress Series*, *436*, 231–243. doi: 10.3354/meps09246

Poulter, M., Hollox, E., Harvey, C. B., Mulcare, C., Peuhkuri, K., Kajander, K., … Swallow, D. M. (2003). The Causal Element for the Lactase Persistence/ non-persistence Polymorphism is Located in a 1 Mb Region of Linkage Disequilibrium in Europeans. *Annals of Human Genetics*, *67*(4), 298–311. doi: 10.1046/j.1469-1809.2003.00048.x

Prakash, S. K., Paylor, R., Jenna, S., Lamarche-Vane, N., Armstrong, D. L., Xu, B., … Zoghbi, H. Y. (2000). Functional analysis of ARHGAP6, a novel GTPase-activating protein for RhoA. *Human Molecular Genetics*, *9*(4), 477–488. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10699171

Prestrud, K. W., Åsbakk, K., Fuglei, E., Mørk, T., Stien, A., Ropstad, E., … Oksanen, A. (2007). Serosurvey for Toxoplasma gondii in arctic foxes and possible sources of infection in the high Arctic of Svalbard. *Veterinary Parasitology*, *150*(1–2), 6–12. doi: 10.1016/j.vetpar.2007.09.006

Puigserver, P., Rhee, J., Donovan, J., Walkey, C. J., Yoon, J. C., Oriente, F., … Spiegelman, B. M. (2003). Insulin-regulated hepatic gluconeogenesis through FOXO1|[ndash]|PGC-1|[alpha]| interaction. *Nature*, *423*(6939), 550–555. doi: 10.1038/nature01606

Punetha, J., Kesari, A., Hoffman, E. P., Gos, M., Kamińska, A., Kostera-Pruszczyk, A., … JĘdrzejowska, M. (2017). Novel Col12A1 variant expands the clinical picture of congenital myopathies with extracellular matrix defects. *Muscle and Nerve*, *55*(2), 277–281. doi: 10.1002/mus.25232

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. doi: 10.1086/519795

Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics*, *13*(1), 70. doi: 10.1186/1471-2156-13-70

Raaschou-Nielsen, O., Andersen, Z. J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., … Hoek, G. (2013). Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *The Lancet Oncology*, *14*(9), 813–822. doi: 10.1016/S1470-2045(13)70279-1

Ramalingam, S. (2012). Reduced Expression of RNA Binding Protein CELF2, a Putative Tumor Suppressor Gene in Colon Cancer. *ImmunoGastroenterology*, *1*(1), 27. doi: 10.7178/ig.1.1.7

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, *17*(1), 285–293. doi: 10.1111/j.1365-294X.2007.03417.x

Randi, E., Alves, P. C., Carranza, J., Milosevic-Zlatanovi, S., Sfougaris, A., & Mucci, N. (2004). Phylogeography of roe deer (Capreolus capreolus) populations: The effects of historical genetic subdivisions and recent nonequilibrium dynamics. *Molecular Ecology*, *13*(10), 3071–3083. doi: 10.1111/j.1365-294X.2004.02279.x

Rebay, I., Silver, S. J., & Tootle, T. L. (2005). New vision from Eyes absent: Transcription factors as enzymes. *Trends in Genetics*, *21*(3), 163–171. doi: 10.1016/j.tig.2005.01.005

Redies, C., Hertel, N., & Hübner, C. A. (2012). Cadherins and neuropsychiatric disorders. *Brain Research*, *1470*, 130–144. doi: 10.1016/j.brainres.2012.06.020

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., & Vilo, J. (2016). g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Research*, *44*(W1), W83–W89. doi: 10.1093/nar/gkw199

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, *35*(suppl_2), W193–W200. doi: 10.1093/nar/gkm226

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. doi: 10.1111/mec.13322

Renaut, S., Nolte, A. W., Rogers, S. M., Derome, N., & Bernatchez, L. (2010). SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (Coregonus spp.). *Molecular Ecology*, *20*(3), 545–559. doi: 10.1111/j.1365-294x.2010.04952.x

Reperant, L. A., Hegglin, D., Fischer, C., Kohler, L., Weber, J. M., & Deplazes, P. (2007). Influence of urbanization on the epidemiology of intestinal helminths of the red fox (*Vulpes vulpes*) in Geneva, Switzerland. *Parasitology Research*, *101*(3), 605–611. doi: 10.1007/s00436-007-0520-0

Ricketts, S. L., Pettitt, L., McLaughlin, B., Jenkins, C. A., & Mellersh, C. S. (2015). A novel locus on canine chromosome 13 is associated with cataract in the Australian Shepherd breed of domestic dog. *Mammalian Genome*, *26*(5–6), 257–263. doi: 10.1007/s00335-015-9562-2

Rocha, N., Payne, F., Huang-doran, I., Sleigh, A., Fawcett, K., Adams, C., … Semple, R. K. (2017). *The metabolic syndrome- associated small G protein ARL15 plays a role in adipocyte differentiation and adiponectin secretion*. (6), 1–12. doi: 10.1038/s41598-017-17746-8

Ronald, J., & Akey, J. M. (2005). Genome-wide scans for loci under selection in humans. *Human Genomics*, *2*(2), 113–125. doi: 10.1186/1479-7364-2-2-113

Rovelli, V., Ruiz-gonzález, A., Vignoli, L., & Macale, D. (2019). Genotyping-by-Sequencing (GBS) of large amphibian genomes : A comparative study of two non-model species endemic to Italy Genotyping-by-Sequencing (GBS) of large amphibian genomes : a comparative study of two non-model species endemic to Italy. *Animal Biology*, (1). doi: 10.1163/15707563-00001094

Rupprecht, C. E., Turmelle, A., & Kuzmin, I. V. (2011a). A perspective on lyssavirus emergence and perpetuation. *Current Opinion in Virology*, *1*(6), 662–670. doi: 10.1016/j.coviro.2011.10.014

Russell, J., Hackett, C., Hedley, P., Liu, H., Milne, L., Bayer, M., … Brennan, R. (2014). The use of genotyping by sequencing in blackcurrant (Ribes nigrum): developing high-resolution linkage maps in species without reference genome sequences. *Molecular Breeding*, *33*(4), 835–849. doi: 10.1007/s11032-013-9996-8

Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., … Lander, E. S. (2006). *Supporting Online Material for Positive Natural Selection in the Human Lineage*. *1641*(6), 1614–1621. doi: 10.1126/science.1124309

Sablin, M., & Germonpre, M. (2004). Systematics and osteometry of Late Glacial foxes from Belgium. *Bulletin de l'institut Royal Des Sciences Naturelles de Belqique, Sciences de La Terre*, (74), 175–188.

Sacks, B. N., Statham, M. J., Perrine, J. D., Wisely, S. M., & Aubry, K. B. (2010). North American montane red foxes: Expansion, fragmentation, and the origin of the Sacramento Valley red fox. *Conservation Genetics*, *11*(4), 1523–1539. doi: 10.1007/s10592-010-0053-4

Santos, J. S., Fonseca, N. A., Vieira, C. P., Vieira, J., & Casares, F. (2010). *Phylogeny of the Teashirt-related Zinc Finger (tshz) Gene Family and Analysis of the Developmental Expression of tshz2 and tshz3b in the Zebrafish*. (1), 1010–1018. doi: 10.1002/dvdy.22228

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807–820. doi: 10.1038/nrg3522

Saxena, N., & Kumar, V. (2014). The HBx Oncoprotein of Hepatitis B Virus Deregulates the Cell Cycle by Promoting the Intracellular Accumulation and Re-Compartmentalization of the Cellular Deubiquitinase USP37. *PLoS ONE*, *9*(10), e111256. doi: 10.1371/journal.pone.0111256

Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. Plant Biotechnology Journal, 15(2), 149-161. doi: 10.1111/pbi.12645

Schmidt, K., Jędrzejewski, W., Theuerkauf, J., Kowalczyk, R., Okarma, H., & Jędrzejewska, B. (2008). Reproductive behaviour of wild-living wolves in Białowieża Primeval Forest (Poland). *Journal of*

Ethology, *26*(1), 69–78. doi: 10.1007/s10164-006-0031-y

Schmidt, P. S., Duvernell, D. D., & Eanes, W. F. (2002). Adaptive evolution of a candidate gene for aging in Drosophila. *Proceedings of the National Academy of Sciences*, *97*(20), 10861–10865. doi: 10.1073/pnas.190338897

Scholz, H. C., Hofer, E., Vergnaud, G., Fleche, P. Le, Whatmore, A. M., Dahouk, S. Al, … Tomaso, H. (2009). Isolation of Brucella microti from Mandibular Lymph Nodes of Red Foxes, *Vulpes vulpes* , in Lower Austria. *Vector-Borne and Zoonotic Diseases*, *9*(2), 153–156. doi: 10.1089/vbz.2008.0036

Schwabe, G. C., Hoffmann, K., Loges, N. T., Birker, D., Rossier, C., de Santi, M. M., … Bartoloni, L. (2008). Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Human Mutation*, *29*(2), 289–298. doi: 10.1002/humu.20656

Schwartz, M., Ralls, K., Williams, D., Cypher, B., Pilgrim, K., & Fleischer, R. (2005). Gene flow among San Joaquin kit fox populations in a severely changed ecosystem. *Conservation Genetics*, *6*(1), 25-37. doi: 10.1007/s10592-004-7719-8

Schweizer, R. M., Robinson, J., Harrigan, R., Silva, P., Galverni, M., Musiani, M., … Wayne, R. K. (2016). Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. *Molecular Ecology*, *25*(1), 357–379. doi: 10.1111/mec.13467

Schweizer, R. M., VonHoldt, B. M., Harrigan, R., Knowles, J. C., Musiani, M., Coltman, D., … Wayne, R. K. (2016). Genetic subdivision and candidate genes under selection in North American grey wolves. *Molecular Ecology*, *25*(1), 380–402. doi: 10.1111/mec.13364

Scott, D. M., Berg, M. J., Tolhurst, B. A., Chauvenet, A. L. M., Smith, G. C., Neaves, K., … Baker, P. J. (2014). Changes in the Distribution of Red Foxes (*Vulpes vulpes*) in Urban Areas in Great Britain: Findings and Limitations of a Media-Driven Nationwide Survey. *PLoS ONE*, *9*(6), e99059. doi: 10.1371/journal.pone.0099059

Selimi, F., Lohof, A. M., Heitz, S., Lalouette, A., Jarvis, C. I., Bailly, Y., & Mariani, J. (2003). Lurcher GRID2-Induced Death and Depolarization Can Be Dissociated in Cerebellar Purkinje Cells. *Neuron*, *37*(5), 813–819. doi: 10.1016/S0896-6273(03)00093-X

Sezgin, E., Duvernell, D. D., Matzkin, L. M., Duan, Y., Zhu, C. T., Verrelli, B. C., & Eanes, W. F. (2004). Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. *Genetics*, *168*(2), 923–931. doi: 10.1534/genetics.104.027649

Sgrò, C. M., Lowe, A. J., & Hoffmann, A. A. (2011). Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, *4*(2), 326–337. doi: 10.1111/j.1752-4571.2010.00157.x

She, R., Chu, J. S.-C., Wang, K., Pei, J., & Chen, N. (2008). genBlastA: Enabling BLAST to identify homologous gene sequences. *Genome Research*, *19*(1), 143–149. doi: 10.1101/gr.082081.108

Sheen, C. R., Jewell, U. R., Morris, C. M., Brennan, S. O., Férec, C., George, P. M., … Chen, J.-M. (2007). Double complex mutations involving F8 and FUNDC2 caused by distinct break-induced replication. *Human Mutation*, *28*(12), 1198–1206. doi: 10.1002/humu.20591

Shi, H., Kichaev, G., & Pasaniuc, B. (2016). Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *The American Journal of Human Genetics*, *99*(1), 139–153. doi: 10.1016/j.ajhg.2016.05.013

Shimoyama, M., De Pons, J., Hayman, G. T., Laulederkind, S. J. F., Liu, W., Nigam, R., … Jacob, H.

(2015). The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. *Nucleic Acids Research*, *43*(D1), D743–D750. doi: 10.1093/nar/gku1026

Shin, Y.-H., Ren, Y., Suzuki, H., Golnoski, K. J., Ahn, H. W., Mico, V., & Rajkovic, A. (2017). Transcription factors SOHLH1 and SOHLH2 coordinate oocyte differentiation without affecting meiosis I. *Journal of Clinical Investigation*, *127*(6), 2106–2117. doi: 10.1172/JCI90281

Sigala, J. L. D. (2004). Activation of Transcription Factor NF- B Requires ELKS, an I B Kinase Regulatory Subunit. *Science*, *304*(5679), 1963–1967. doi: 10.1126/science.1098387

Skog, A., Zachos, F. E., Rueness, E. K., Feulner, P. G. D., Mysterud, A., Langvatn, R., … Jakobsen, K. S. (2009). Phylogeography of red deer (Cervus elaphus) in Europe. *Journal of Biogeography*, *36*(1), 66–77. doi: 10.1111/j.1365-2699.2008.01986.x

Smith, A. J., Mondain-Monval, M., MOller, O. M., Scholler, R., & Hansson, V. (1985). Seasonal variations of LH, prolactin, androstenedione, testosterone and testicular FSH binding in the male blue fox (*Alopex lagopus*). *Reproduction*, *74*(2), 449–458. doi: 10.1530/jrf.0.0740449

Smith, G. C., Gangadharan, B., Taylor, Z., Laurenson, M. K., Bradshaw, H., Hide, G., … Craig, P. S. (2003). Prevalence of zoonotic important parasites in the red fox (*Vulpes vulpes*) in Great Britain. *Veterinary Parasitology*, *118*(1–2), 133–142. doi: 10.1016/j.vetpar.2003.09.017

Sommer, R., & Benecke, N. (2005). Late-Pleistocene and early Holocene history of the canid fauna of Europe (Canidae). *Mammalian Biology*, *70*(4), 227–241. doi: 10.1016/j.mambio.2004.12.001

Son, H. Y., Sohn, S. W., Im, S. H., Kim, H. J., Lee, M. K., Gombojav, B., … Kim, J. Il. (2015). Family-based association study of pulmonary function in a population in northeast Asia. *PLoS ONE*, *10*(10), 1–11. doi: 10.1371/journal.pone.0139716

Spurgin, L. G., & Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1684), 979–988. doi: 10.1098/rspb.2009.2084

Sreedhar, A., Petruska, P., Miriyala, S., Panchatcharam, M., & Zhao, Y. (2017). UCP2 overexpression enhanced glycolysis via activation of PFKFB2 during skin cell transformation. *Oncotarget*, *8*(56), 95504–95515. doi: 10.18632/oncotarget.20762

Statham, M. J., Edwards, C. J., Norén, K., Soulsbury, C. D., & Sacks, B. N. (2018). Genetic analysis of European red foxes reveals multiple distinct peripheral populations and central continental admixture. *Quaternary Science Reviews*, *197*, 257–266. doi: 10.1016/j.quascirev.2018.08.019

Statham, Mark J., Murdoch, J., Janecka, J., Aubry, K. B., Edwards, C. J., Soulsbury, C. D., … Sacks, B. N. (2014). Range-wide multilocus phylogeography of the red fox reveals ancient continental divergence, minimal genomic exchange and distinct demographic histories. *Molecular Ecology*, *23*(19), 4813–4830. doi: 10.1111/mec.12898

Steck, F., & Wandeler, A. (1980a). The epidemiology of fox rabies in Europe. *Epidemiologic Reviews*, *2*, 71–96.

Steck, F., & Wandeler, A. (1980b). The epidemiology of fox rabies in Europe. *Epidemiologic Reviews*, *2*, 71–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7000539

Stembridge, M., Ainslie, P. N., Hughes, M. G., Stöhr, E. J., Cotter, J. D., Nio, A. Q. X., & Shave, R. (2014). Ventricular structure, function, and mechanics at high altitude: chronic remodeling in Sherpa vs. short-term lowlander adaptation. *Journal of Applied Physiology*, *117*(3), 334–343. doi: 10.1152/japplphysiol.00233.2014

Svendsen, J. M., Smogorzewska, A., Sowa, M. E., O'Connell, B. C., Gygi, S. P., Elledge, S. J., & Harper,

J. W. (2009). Mammalian BTBD12/SLX4 Assembles A Holliday Junction Resolvase and Is Required for DNA Repair. *Cell*, *138*(1), 63–77. doi: 10.1016/j.cell.2009.06.030

Swanson, B. J., Fuhrmann, R. T., & Crabtree, R. L. (2005). Elevational isolation of red fox populations in the Greater Yellowstone Ecosystem. *Conservation Genetics*, *6*(1), 123–131. doi: 10.1007/s10592-004-7730-0

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*(4), 453–464. doi: 10.1046/j.1365-294x.1998.00289.x

Team, R. (2016). RStudio: integrated development for R.(RStudio, Inc., Boston, MA, USA).

Takayanagi-Kiya, S., Kiya, T., Kunieda, T., & Kubo, T. (2017). Mblk-1 Transcription Factor Family: Its Roles in Various Animals and Regulation by NOL4 Splice Variants in Mammals. *International Journal of Molecular Sciences*, *18*(2), 246. doi: 10.3390/ijms18020246

Tang, Y., Horikoshi, M., & Li, W. (2016). ggfortify: Unified Interface to Visualize Statistical Results of Popular R Packages. *The R Journal*, *8*(2), 474. doi: 10.32614/RJ-2016-060

Teacher, A. G., Thomas, J. A., & Barnes, I. (2011). Modern and ancient red fox (Vulpes vulpes) in Europe show an unusual lack of geographical and temporal structuring, and differing responses within the carnivores to historical climatic change. BMC Evolutionary Biology, 11(1). doi: 10.1186/1471-2148-11-214

Tetteh, K. K. A., Stewart, L. B., Ochola, L. I., Amambua-Ngwa, A., Thomas, A. W., Marsh, K., … Conway, D. J. (2009). Prospective identification of malaria parasite genes under balancing selection. *PLoS ONE*, *4*(5). doi: 10.1371/journal.pone.0005568

Tews, D., Fromme, T., Keuper, M., Hofmann, S. M., Debatin, K. M., Klingenspor, M., … Fischer-Posovszky, P. (2017). Teneurin-2 (TENM2) deficiency induces UCP1 expression in differentiating human fat cells. *Molecular and Cellular Endocrinology*, *443*, 106–113. doi: 10.1016/j.mce.2017.01.015

Thomas, C., Rousset, R., & Noselli, S. (2009). JNK signalling influences intracellular trafficking during Drosophila morphogenesis through regulation of the novel target gene Rab30. *Developmental Biology*, *331*(2), 250–260. doi: 10.1016/j.ydbio.2009.05.001

Thomas, L., Kennington, W. J., Evans, R. D., Kendrick, G. A., & Stat, M. (2017). Restricted gene flow and local adaptation highlight the vulnerability of high-latitude reefs to rapid environmental change. *Global Change Biology*, *23*(6), 2197–2205. doi: 10.1111/gcb.13639

Tiwary, B. K. (2016). Evolution of the *SRGAP2* Gene Is Linked to Intelligence in Mammals. *Biomedicine Hub*, *1*(1), 443947–443947. doi: 10.1159/000443947

Tokumasu, Y., Iida, A., Wang, Z., Ansai, S., Kinoshita, M., & Sehara-Fujisawa, A. (2016). ADAM12-deficient zebrafish exhibit retardation in body growth at the juvenile stage without developmental defects. *Development Growth and Differentiation*, *58*(4), 409–421. doi: 10.1111/dgd.12286

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. doi: 10.1037/h0061626

Toyofuku, T., Zhang, H., Kumanogoh, A., Takegahara, N., Yabuki, M., Harada, K., … Kikutani, H. (2004). Guidance of myocardial patterning in cardiac development by Sema6D reverse signalling. *Nature Cell Biology*, *6*(12), 1204–1211. doi: 10.1038/ncb1193

Trouwborst, A. (2014). Exploring the Legal Status of Wolf-Dog Hybrids and Other Dubious Animals: International and EU Law and the Wildlife Conservation Problem of Hybridization with Domestic and

Alien Species. *Review of European, Comparative & International Environmental Law*, *23*(1), 111–124. doi: 10.1111/reel.12052

Tucker, A., & Sharpe, P. (2004). The cutting-edge of mammalian development; how the embryo makes teeth. *Nature Reviews Genetics*, *5*(7), 499–508. doi: 10.1038/nrg1380

Uchida, Y. (2014). Ceramide signaling in mammalian epidermis. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1841*(3), 453–462. doi: 10.1016/j.bbalip.2013.09.003

Ullah, I., Kakar, N., Schrauwen, I., Hussain, S., Chakchouk, I., Liaqat, K., … Leal, S. M. (2019). Variants in KIAA0825 underlie autosomal recessive postaxial polydactyly. *Human Genetics*, *0*(0), 1–8. doi: 10.1007/s00439-019-02000-0

Vagnozzi, R. J., Gatto, G. J., Kallander, L. S., Hoffman, N. E., Mallilankaraman, K., Ballard, V. L. T., … Force, T. (2013). Inhibition of the Cardiomyocyte-Specific Kinase TNNI3K Limits Oxidative Stress, Injury, and Adverse Remodeling in the Ischemic Heart. *Science Translational Medicine*, *5*(207), 207ra141-207ra141. doi: 10.1126/scitranslmed.3006479

Valdespino, C., Asa, C. S., & Bauman, J. E. (2005). Estrous Cycles, Copulation, and Pregnancy in the Fennec Fox (Vulpes Zerda). *Journal of Mammalogy*, *83*(1), 99–109. doi: 10.1644/1545-1542(2002)083<0099:eccapi>2.0.co;2

Valentino, E., Bellazzo, A., Di Minin, G., Sicari, D., Apollonio, M., Scognamiglio, G., … Collavin, L. (2017). Mutant p53 potentiates the oncogenic effects of insulin by inhibiting the tumor suppressor DAB2IP. *Proceedings of the National Academy of Sciences*, *114*(29), 7623–7628. doi: 10.1073/pnas.1700996114

Valero-Rubio, D., Jiménez, K. M., Fonseca, D. J., Payán-Gómez, C., & Laissue, P. (2018). Transcriptomic analysis of FUCA1 knock-down in keratinocytes reveals new insights into the pathogenesis of fucosidosis skin lesions. *Experimental Dermatology*, *27*(6), 663–667. doi: 10.1111/exd.13532

Vasco, V. R. Lo, Cardinale, G., & Polonia, P. (2012). Deletion of PLCB1 gene in schizophrenia-affected patients. *Journal of Cellular and Molecular Medicine*, *16*(4), 844–851. doi: 10.1111/j.1582-4934.2011.01363.x

Vecoli, C., Montano, L., & Andreassi, M. G. (2016). Environmental pollutants: genetic damage and epigenetic changes in male germ cells. *Environmental Science and Pollution Research*, *23*(23), 23339–23348. doi: 10.1007/s11356-016-7728-4

Veeriah, S., Brennan, C., Meng, S., Singh, B., Fagin, J. A., Solit, D. B., … Chan, T. A. (2009). The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. *Proceedings of the National Academy of Sciences*, *106*(23), 9435–9440. doi: 10.1073/pnas.0900571106

Veness, C. (2019). Calculate distance and bearing between two Latitude/Longitude points using haversine formula in JavaScript. [online] Movable-type.co.uk. Retrieved from www.movable-type.co.uk/scripts/latlong.html [Accessed 2 Feb. 2018].

Verhagen, A. M., Ekert, P. G., Pakusch, M., Silke, J., Connolly, L. M., Reid, G. E., … Vaux, D. L. (2000). Identification of DIABLO, a Mammalian Protein that Promotes Apoptosis by Binding to and Antagonizing IAP Proteins. *Cell*, *102*(1), 43–53. doi: 10.1016/S0092-8674(00)00009-X

Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement The Kappa Statistic. *Fam Med*, *37*(5), 360–363.

Virgós, E., Kowalczyk, R., Trua, A., de Marinis, A., Mangas, J. G., Barea-Azcón, J. M., & Geffen, E.

(2011). Body size clines in the European badger and the abundant centre hypothesis. *Journal of Biogeography*, *38*(8), 1546–1556. doi: 10.1111/j.1365-2699.2011.02512.x

Volcik, K. A., Catellier, D., Folsom, A. R., Matijevic, N., Wasserman, B., & Boerwinkle, E. (2009). SELP and SELPLG Genetic Variation Is Associated with Cell Surface Measures of SELP and SELPLG: The Atherosclerosis Risk in Communities Carotid MRI Study. *Clinical Chemistry*, *55*(6), 1076–1082. doi: 10.1373/clinchem.2008.119487

VonHoldt, B. M., Cahill, J. A., Fan, Z., Gronau, I., Robinson, J., Pollinger, J. P., … Wayne, R. K. (2016). Whole-genome sequence analysis shows two endemic species of North American wolf are admixtures of the coyote and gray wolf. *Science Advances*, *2*(7). doi: 10.1126/sciadv.1602250

Vuorisalo, T., Talvitie, K., Kauhala, K., Bläuer, A., & Lahtinen, R. (2014). Urban red foxes (*Vulpes vulpes L*.) in Finland: A historical perspective. *Landscape and Urban Planning*, *124*, 109–117. doi: 10.1016/j.landurbplan.2013.12.002

Waardenberg, A. J., Bernardo, B. C., Ng, D. C. H., Shepherd, P. R., Cemerlang, N., Sbroggiò, M., … McMullen, J. R. (2011). Phosphoinositide 3-Kinase (PI3K(p110α)) Directly Regulates Key Components of the Z-disc and Cardiac Structure. *Journal of Biological Chemistry*, *286*(35), 30837–30846. doi: 10.1074/jbc.M111.271684

Walton, Z., Samelius, G., Odden, M., & Willebrand, T. (2017). Variation in home range size of red foxes *Vulpes vulpes* along a gradient of productivity and human landscape alteration. *PLoS ONE*, *12*(4), 1–14. doi: 10.1371/journal.pone.0175291

Walton, Z., Samelius, G., Odden, M., & Willebrand, T. (2018). Long-distance dispersal in red foxes *Vulpes vulpes* revealed by GPS tracking. *European Journal of Wildlife Research*, *64*(6), 64. doi: 10.1007/s10344-018-1223-9

Wang, J., Lv, X., Xu, F., Wei, M., Liu, C., & Yang, Y. (2018). GNA14 silencing suppresses the proliferation of endometrial carcinoma cells through inducing apoptosis and G 2 /M cell cycle arrest. *Bioscience Reports*, *38*(5), BSR20180574. doi: 10.1042/BSR20180574

Wang, L., Feng, Y., Yan, D., Qin, L., Grati, M., Mittal, R., … Liu, X. (2018). A dominant variant in the PDE1C gene is associated with nonsyndromic hearing loss. *Human Genetics*, *137*(6–7), 437–446. doi: 10.1007/s00439-018-1895-y

Wang, Y., Wang, S., Wang, L., Yao, Y., Ma, C., Ding, J., … Li, J. (2017). Overexpression of Cardiac-Specific Kinase TNNI3K Promotes Mouse Embryonic Stem Cells Differentiation into Cardiomyocytes. *Cellular Physiology and Biochemistry*, *41*(1), 381–398. doi: 10.1159/000456400

Wayne, R. K. (1993). Molecular evolution of the dog family. *Trends in Genetics*, *9*(6), 218–224. Retrieved from http://www.sciencedirect.com/science/article/pii/016895259390122X

Webster, J. P., Lamberton, P. H. L., Donnelly, C. A., & Torrey, E. F. (2006). Parasites as causative agents of human affective disorders? The impact of anti-psychotic, mood-stabilizer and anti-parasite medication on Toxoplasma gondii's ability to alter host behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1589), 1023–1030. doi: 10.1098/rspb.2005.3413

Weedall, G. D., & Conway, D. J. (2010). Detecting signatures of balancing selection to identify targets of anti-parasite immunity. *Trends in Parasitology*, *26*(7), 363–369. doi: 10.1016/j.pt.2010.04.002

Weir, B. S., & Cockerham, C. C. (1984). *Atistics for the analysis of population structure*. *38*(6), 1358–1370.

Weitnauer, M., Mijošek, V., & Dalpke, A. H. (2016). Control of local immunity by airway epithelial cells. *Mucosal Immunology*, *9*(2), 287–298. doi: 10.1038/mi.2015.126

Wells, M. C., & Lehner, P. N. (1978). The relative importance of the distance senses in coyote predatory behaviour. *Animal Behaviour*, *26*(01), 251–258. doi: 10.1016/0003-3472(78)90025-8

Wheat, C. W. (2010). Rapidly developing functional genomics in ecological model systems via 454 transcriptome sequencing. *Genetica*, *138*(4), 433–451. doi: 10.1007/s10709-008-9326-y

Wiezlak, M., Diring, J., Abella, J., Mouilleron, S., Way, M., McDonald, N. Q., & Treisman, R. (2012). G-actin regulates the shuttling and PP1 binding of the RPEL protein Phactr1 to control actomyosin assembly. *Journal of Cell Science*, *125*(23), 5860–5872. doi: 10.1242/jcs.112078

Wilkins, M. R., Karaardıç, H., Vortman, Y., Parchman, T. L., Albrecht, T., Petrželková, A., … Safran, R. J. (2016). Phenotypic differentiation is associated with divergent sexual selection among closely related barn swallow populations. *Journal of Evolutionary Biology*, *29*(12), 2410–2421. doi: 10.1111/jeb.12965

Williams, J. B., Muñoz-Garcia, A., Ostrowski, S., & Tieleman, B. I. (2004). A phylogenetic analysis of basal metabolism, total evaporative water loss, and life-history among foxes from desert and mesic regions. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *174*(1), 29–39. doi: 10.1007/s00360-003-0386-0

Wood, J. (1958). Age Structure and Productivity of a Gray Fox Population. *Journal Of Mammalogy*, *39*(1), 74. doi: 10.2307/1376612

Wu, T., Wang, X., Wei, C., Cheng, H., Wang, X., & Li, Y. et al. (2005). Hemoglobin levels in Qinghai-Tibet: different effects of gender for Tibetans vs. Han. Journal Of Applied Physiology, 98(2), 598-604. doi: 10.1152/japplphysiol.01034.2002

Xu, Y., Xue, S., Zhou, J., Voorhees, J. J., & Fisher, G. J. (2015). Notch and TGF-β pathways cooperatively regulate receptor protein tyrosine phosphatase-κ (PTPRK) gene expression in human primary keratinocytes. *Molecular Biology of the Cell*, *26*(6), 1199–1206. doi: 10.1091/mbc.E14-12-1591

Yagi, Y., & Ip, Y. T. (2005). Helicase89B is a Mot1p/BTAF1 homologue that mediates an antimicrobial response in Drosophila. *EMBO Reports*, *6*(11), 1088–1094. doi: 10.1038/sj.embor.7400542

Yamada, N., Asano, Y., Fujita, M., Yamazaki, S., Inanobe, A., Matsuura, N., … Takashima, S. (2019). Mutant KCNJ3 and KCNJ5 Potassium Channels as Novel Molecular Targets in Bradyarrhythmias and Atrial Fibrillation. *Circulation*, *139*(18), 2157–2169. doi: 10.1161/CIRCULATIONAHA.118.036761

Yan, W. (2009). Male infertility caused by spermiogenic defects: Lessons from gene knockouts. *Molecular and Cellular Endocrinology*, *306*(1–2), 24–32. doi: 10.1016/j.mce.2009.03.003

Yang, Y., Lu, Y., Espejo, A., Wu, J., Xu, W., Liang, S., & Bedford, M. T. (2010). TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks. *Molecular Cell*, *40*(6), 1016–1023. doi: 10.1016/j.molcel.2010.11.024

Yasumura, M., Yoshida, T., Yamazaki, M., Abe, M., Natsume, R., Kanno, K., … Mishina, M. (2015). IL1RAPL1 knockout mice show spine density decrease, learning deficiency, hyperactivity and reduced anxiety-like behaviours. *Scientific Reports*, *4*(1), 6613. doi: 10.1038/srep06613

Ye, L., Lin, S. tao, Mi, Y. shuai, Liu, Y., Ma, Y., Sun, H. min, … Fan, J. wei. (2016). Overexpression of LARP1 predicts poor prognosis of colorectal cancer and is expected to be a potential therapeutic target. *Tumor Biology*, *37*(11), 14585–14594. doi: 10.1007/s13277-016-5332-3

Yeşil, G., Aralaşmak, A., Akyüz, E., İçağasıoğlu, D., Şahin, T. U., & Bayram, Y. (2018). Expanding the phenotype of homozygous kcnma1 mutations; dyskinesia, epilepsy, intellectual disability, cerebellar

and corticospinal tract atrophy. *Balkan Medical Journal*, *35*(4), 336–339. doi: 10.4274/balkanmedj.2017.0986

Yokota, M., Masaki, H., Okano, Y., & Tokudome, Y. (2017). Effect of glycation focusing on the process of epidermal lipid synthesis in a reconstructed skin model and membrane fluidity of stratum corneum lipids. *Dermato-Endocrinology*, *9*(1), e1338992. doi: 10.1080/19381980.2017.1338992

Yom-Tov, Y., Benjamini, Y., & Kark, S. (2002). Global warming, Bergmann's rule and body mass - Are they related? The chukar partridge (Alectoris chukar) case. *Journal of Zoology*, *257*(4), 449–455. doi: 10.1017/S095283690200105X

Zaki, T., & Choate, K. (2018). Recent advances in understanding inherited disorders of keratinization. *F1000Research*, *7*(0), 919. doi: 10.12688/f1000research.14514.1

Zarkower, D., Sarver, A. L., Bardwell, V. J., Griswold, M. D., Murphy, M. W., & Matson, C. K. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature*, *476*(7358), 101–104. doi: 10.1038/nature10239

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., … Linnarsson, S. (2018). Molecular Architecture of the Mouse Nervous System. *Cell*, *174*(4), 999-1014.e22. doi: 10.1016/j.cell.2018.06.021

Zhan, S., Zhang, W., Niitepõld, K., Hsu, J., Haeger, J. F., Zalucki, M. P., … Kronforst, M. R. (2014). The genetics of monarch butterfly migration and warning colouration. *Nature*, *514*(7522), 317–321. doi: 10.1038/nature13812

Zhang, C., Kho, Y. S., Wang, Z., Chiang, Y. T., Ng, G. K. H., Shaw, P. C., … Qi, R. Z. (2014). Transmembrane and coiled-coil domain family 1 is a novel protein of the endoplasmic reticulum. *PLoS ONE*, *9*(1). doi: 10.1371/journal.pone.0085206

Zhang, D., Sliwkowski, M. X., Mark, M., Frantz, G., Akita, R., Sun, Y., … Godowski, P. J. (1997). Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(18), 9562–9567. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9275162%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC23218

Zhang, Q., Liu, J.-F., Wang, J., Wang, H., Wang, Z., Kang, H., … Yin, Z. (2016). Structural Variant Detection by Large-scale Sequencing Reveals New Evolutionary Evidence on Breed Divergence between Chinese and European Pigs. *Scientific Reports*, *6*(1), 1–15. doi: 10.1038/srep18501

Zhang, Wei, Thevapriya, S., Kim, P. J., Yu, W. P., Je, H. S., Tan, E. K. in., & Zeng, L. (2014). Amyloid precursor protein regulates neurogenesis by antagonizing miR-574-5p in the developing cerebral cortex. *Nature Communications*, *5*, 3330. doi: 10.1038/ncomms4330

Zhang, Wenping, Fan, Z., Han, E., Hou, R., Zhang, L., Galaverni, M., … Zhang, Z. (2014). Hypoxia Adaptations in the Grey Wolf (Canis lupus chanco) from Qinghai-Tibet Plateau. *PLoS Genetics*, *10*(7). doi: 10.1371/journal.pgen.1004466

Zhang, H., Wu, C., Chamba, Y., & Ling, Y. (2007). Blood Characteristics for High Altitude Adaptation in Tibetan Chickens. *Poultry Science*, *86*(7), 1384-1389. doi: 10.1093/ps/86.7.1384

Zhang, Y., Stehling-Sun, S., Lezon-Geyda, K., Juneja, S. C., Coillard, L., Chatterjee, G., … Perkins, A. S. (2011). PR-domain - containing Mds1-Evi1 is critical for long-term hematopoietic stem cell function. *Blood*, *118*(14), 3853–3861. doi: 10.1182/blood-2011-02-334680

Zhao, L., Svingen, T., Ng, E. T., & Koopman, P. (2015). Female-to-male sex reversal in mice caused by

transgenic overexpression of Dmrt1. *Development*, *142*(6), 1083–1088. doi: 10.1242/dev.122184

Zhou, D., Udpa, N., Ronen, R., Stobdan, T., Liang, J., Appenzeller, O., … Haddad, G. G. (2013). Whole-Genome sequencing uncovers the genetic basis of chronic mountain sickness in andean highlanders. *American Journal of Human Genetics*, *93*(3), 452–462. doi: 10.1016/j.ajhg.2013.07.011

Zimen, E. (1984). Long range movements of the red fox *Vulpes vulpes* L. *Acta Zoologica Fennica*, *171*(1), 267–270.

Zou, D., Erickson, C., Kim, E.-H., Jin, D., Fritzsch, B., & Xu, P.-X. (2008). Eya1 gene dosage critically affects the development of sensory epithelia in the mammalian inner ear. *Human Molecular Genetics*, *17*(21), 3340–3356. doi: 10.1093/hmg/ddn229

Zwick, W. R., & Velicer, W. F. (1982). Factors Influencing Four Rules For Determining The Number Of Components To Retain. *Multivariate Behavioral Research*, *17*(2), 253–269. doi: 10.1207/s15327906mbr1702_5

Appendices

Appendix I.

Outliers that tested for positive diversifying selection and balancing selection using the BayeScan method. Linked at divergent and balancing selection bayescan 1556 K3 12-05-2019 Appendix I.xlsx (see excel).

Appendix II.

Candidate genes that was found to be uncharacterised.

Gene	
LOC102153862	Uncharacterized
LOC10251436	N/A
LOC106559409	Uncharacterized
LOC106559439	Uncharacterized
LOC111093380	Uncharacterized
LOC111094857	Uncharacterized
LOC111098341	Uncharacterized
LOC111098526	Uncharacterized
LOC478449	Uncharacterized
LOC487987	N/A
LOC100687178	Uncharacterized
ROB2	N/A

Appendix III.

Details of each candidate gene found including SNPs, full gene name, reference and alternate allele (major and minor allele), E-value from BLAST software, whether the SNP occurred in exon or intron, FST, alpha-values, and other information from BayeScan, P-value from PCAdapt, and strand of nucleotides surrounding target SNP from SAM file.
K3_everything_161SNPS_for_Appendix III 12-05-2019.xlsx (see excel).

Appendix IV.

Results from candidate gene list used in g:Profiler with enrichment test showing terms with p-values of <0.05 after correction with multiple testing using the Benjamini-Hochberg FDR method. Linked at gProfiler_cfamiliaris_4-10-2019_7-15-23 PM__intersections K3 12-05-2019 Appendix IV.csv (see excel).

Appendix V.

Results from Ease score of >1.3 (P-value <0.05) the Functional Annotational Clustering candidate gene list used in DAVID. Linked at 38 clusters 0.05 E_score Appendix V.xlsx (see excel).

Appendix VI.

Candidate genes list in each term from DAVID (see Appendix V).

Category	Term	Genes	P-Value	Benjamini
GO	actin filament bundle assembly	ARHGAP6, ARRB1, PHACTR1, PCDH15	2.7E-2	6.1E-1
GO	actin filament bundle organization	ARHGAP6, ARRB1, PHACTR1, PCDH15	2.9E-2	6.2E-1
GO	adult walking behavior	DAB1, KCNMA1, PCDH15	1.8E-2	5.4E-1
GO	anatomical structure development	AFF3, AGFG1, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SMYD3, SRGAP2, SSX2IP, VANGL1, ANK2, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	1.2E-4	1.3E-1
GO	anatomical structure morphogenesis	AFF3, AGFG1, DAB1, DAB2IP, DSCAML1, EYA1, SRGAP2, SSX2IP, CTNND2, COL12A1, DMRT1, EDA, EXOC4, FOXP2, LRRC4C, MAP2, PTCH1, PCSK5, PCDH15, ROBO2, RYR2, SEMA6D, SULF1, TENM2, THRB	2.4E-3	2.7E-1
GO	animal organ development	DAB1, DAB2IP, DSCAML1, EYA1, PPARGC1A, SRGAP2, ANK2, CERS3, CDK6, DMRT1, DNAH11, EDA, EXOC4, FOXP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, THRB	2.9E-3	3.1E-1
GO	axon part	DAB2IP, DLG2, KCNMA1, ROBO2	2.6E-2	4.8E-1
GO	biological adhesion	DAB1, DSCAML1, EDIL3, ARHGAP6, CDH18, CTNND2, CNTN6, CDK6, EDA, PTPRK, PCDH15, ROBO2, SELP, SULF1, TENM2	3.8E-3	3.1E-1
GO	blood circulation	TNNI3K, ANK2, KCNMA1, PCSK5, RYR2, THRB	4.6E-2	7.0E-1
GO	brain development	DAB1, DAB2IP, DSCAML1, SRGAP2, FOXP2, PTCH1, PLCB1, ROBO2, SEMA6D	1.1E-2	4.6E-1

GO	cell adhesion	DAB1, DSCAML1, EDIL3, ARHGAP6, CDH18, CTNND2, CNTN6, CDK6, EDA, PTPRK, PCDH15, ROBO2, SELP, SULF1, TENM2	3.5E-3	3.2E-1
GO	cell communication	ARL15, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, MECOM, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SRGAP2, SSX2IP, ANK2, ARRB1, CTNND2, CDK6, DIABLO, DLG2, DMRT1, EDA, FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, MYH13, PTCH1, PIK3CA, PDE1C, PLCB1, KCNMA1, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	5.0E-3	3.4E-1
GO	cell development	AGFG1, DAB1, DAB2IP, EYA1, SMYD3, SRGAP2, CTNND2, CDK6, DMRT1, LRRC4C, MAP2, KCNMA1, PCDH15, ROBO2, SEMA6D, SOHLH2, SULF1, TENM2	1.6E-2	5.4E-1
GO	cell differentiation	AGFG1, DAB1, DAB2IP, EYA1, SMYD3, SRGAP2, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DMRT1, FOXO1, LRRC4C, MAP2, PTCH1, PLCB1, KCNMA1, POMGNT2, PCDH15, RIPK1, ROBO2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	8.6E-3	4.0E-1
GO	cell morphogenesis	DAB1, DAB2IP, SRGAP2, SSX2IP, CTNND2, DMRT1, LRRC4C, MAP2, PCDH15, ROBO2, SEMA6D, TENM2	4.9E-2	7.00E-01
GO	cell part morphogenesis	DAB1, DAB2IP, SRGAP2, SSX2IP, CTNND2, MAP2, PCDH15, ROBO2, SEMA6D	4.90E-02	7.20E-01
GO	cell periphery	ADAM12, ABCC1, DAB2IP, NDC1, SRGAP2, VANGL1, ANK2, ANO2, CDH18, DIABLO, DAGLB, DLG2, EDA, EXOC4, GABRB1, GRID2, GUCY1A2, IL1RAPL1, NBEA, PARD6B, PTCH1, PIK3CA, KCNMA1, KCNJ3, PTPRK, PCDH15, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	1.3E-2	3.4E-1
GO	cell projection	AGFG1, DAB2IP, SSX2IP, ANO2, ARRB1, CDK6, DLG2, DNAH11, EXOC4, MAP2, PTCH1, PIK3CA, PIP5K1B, PDE1C, KCNMA1, PTPRK, PCDH15, ROBO2, TENM2	1.5E-4	3.6E-2
GO	cell projection morphogenesis	DAB1, DAB2IP, SRGAP2, SSX2IP, CTNND2, MAP2, PCDH15, ROBO2, SEMA6D	4.7E-2	7.0E-1

GO	cell-cell adhesion	DAB1, DSCAML1, CDH18, CTNND2, CDK6, PCDH15, ROBO2, SELP, TENM2	4.9E-2	7.6E-1
GO	cell-cell adhesion via plasma-membrane adhesion molecules	DAB1, DSCAML1, CDH18, PCDH15, ROBO2, SELP, TENM2	2.2E-4	9.7E-2
GO	cell-cell signaling	DAB2IP, ANK2, ARRB1, CTNND2, DLG2, EDA, FOXO1, KCNMA1, RYR2, SULF1, TACR1	4.9E-2	7.6E-1
GO	cell-cell signaling by wnt	DAB2IP, CTNND2, EDA, FOXO1, RYR2, SULF1	3.5E-2	6.6E-1
GO	cellular component morphogenesis	AGFG1, DAB1, DAB2IP, SRGAP2 , SSX2IP, CTNND2, DMRT1, LRRC4C, MAP2, PCDH15, ROBO2, SEMA6D, TENM2	3.5E-2	6.5E-1
GO	cellular developmental process	AGFG1, DAB1, DAB2IP, EYA1, SMYD3, SRGAP2, SSX2IP, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DMRT1, FOXO1, LRRC4C, MAP2, PTCH1, PLCB1, KCNMA1, POMGNT2, PCDH15, RIPK1, ROBO2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	1.7E-2	5.3E-1
GO	cellular response to oxygen-containing compound	DAB2IP, SMYD3, ARRB1, FOXO1, PTCH1, PIK3CA, PLCB1, PTPRK, RYR2	2.3E-2	5.9E-1
GO	cellular response to stimulus	ARL15, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, LARP1B, MECOM, PPARGC1A, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SMYD3, SRGAP2, SSX2IP, ANK2, ARRB1, CTNND2, CDK6, DIABLO, DMRT1, EDA, FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, MYH13, PTCH1, PIK3CA, PDE1C, PLCB1, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	2.9E-2	6.1E-1
GO	central nervous system development	DAB1, DAB2IP, DSCAML1, SRGAP2, FOXP2, MAP2, PTCH1, PLCB1, ROBO2, SEMA6D	1.8E-2	5.4E-1
GO	cerebral cortex cell migration	DAB1, DAB2IP, SRGAP2	3.3E-2	6.4E-1
GO	cerebral cortex development	DAB1, DAB2IP, SRGAP2, FOXP2, PLCB1	2.3E-3	2.8E-1
GO	cerebral cortex radial glia guided migration	DAB1, DAB2IP, SRGAP2	6.9E-3	3.7E-1

GO	cerebral cortex radially oriented cell migration	DAB1, DAB2IP, SRGAP2	1.5E-2	5.3E-1
GO	channel activity	ANO2, GABRB1, GRID2, MTMR6, KCNMA1, KCNJ3, RYR2	4.9E-2	7.1E-1
GO	cilium	SSX2IP, ANO2, DNAH11, PTCH1, PDE1C, PTPRK, PCDH15	2.3E-2	4.8E-1
GO	circulatory system process	TNNI3K, ANK2, KCNMA1, PCSK5, RYR2, THRB	4.8E-2	7.0E-1
GO	contractile actin filament bundle assembly	ARHGAP6, ARRB1, PHACTR1	4.6E-2	7.0E-1
GO	dendrite development	DAB1, DAB2IP, SRGAP2, CTNND2, MAP2	1.2E-2	4.7E-1
GO	developmental process	AFF3, AGFG1, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SMYD3, SRGAP2, SSX2IP, VANGL1, ANK2, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	2.3E-4	8.2E-2
GO	embryo development	AFF3, DSCAML1, EYA1, KIAA1217, COL12A1, EXOC4, PTCH1, PLCB1, PCSK5, PCDH15, RYR2, SULF1	1.7E-2	5.4E-1
GO	embryonic morphogenesis	AFF3, DSCAML1, EYA1, COL12A1, EXOC4, PTCH1, PCDH15, RYR2, SULF1	1.6E-2	5.4E-1
GO	embryonic skeletal system development	DSCAML1, EYA1, KIAA1217, PCSK5, SULF1	7.1E-3	3.7E-1
GO	enzyme regulator activity	AGFG1, AGAP1, DAB2IP, GARNL3, ARHGAP20, ARHGAP6, SRGAP2, ARRB1, LRRC4C, PHACTR1, PHACTR2, PIK3CA, PLCB1	1.6E-3	1.0E-1
GO	epidermal cell differentiation	CERS3, PTCH1, KCNMA1, PCDH15	4.6E-2	7.0E-1
GO	epidermis development	CERS3, EDA, PTCH1, KCNMA1, PCDH15	4.0E-2	6.8E-1
GO	epithelial cell differentiation	EYA1, CERS3, CDK6, DMRT1, PTCH1, KCNMA1, PCDH15, THRB	2.2E-2	5.7E-1
GO	epithelial cell proliferation	DAB2IP, EYA1, CDK6, FOXP2, PTCH1, PTPRK, SULF1	7.6E-3	3.8E-1
GO	epithelium development	EYA1, CERS3, CDK6, DMRT1, EDA, FOXP2, PTCH1, KCNMA1, PCDH15, ROBO2, RYR2, SULF1, THRB	6.2E-3	3.6E-1

GO	gated channel activity	ANO2, GABRB1, GRID2, MTMR6, KCNMA1, KCNJ3, RYR2	1.1E-2	4.1E-1
GO	generation of neurons	DAB1, DAB2IP, EYA1, SRGAP2, CTNND2, CNTN6, DAGLB, LRRC4C, MAP2, PTCH1, KCNMA1, POMGNT2, PCDH15, ROBO2, SEMA6D, TENM2	1.7E-3	2.5E-1
GO	glial cell migration	DAB1, DAB2IP, SRGAP2	1.7E-2	5.4E-1
GO	GTPase regulator activity	AGFG1, AGAP1, DAB2IP, GARNL3, ARHGAP20, ARHGAP6, SRGAP2, ARRB1, PLCB1	5.2E-5	8.8E-3
GO	head development	DAB1, DAB2IP, DSCAML1, SRGAP2, FOXP2, PTCH1, PLCB1, ROBO2, SEMA6D	1.7E-2	5.4E-1
GO	homophilic cell adhesion via plasma membrane adhesion molecules	DSCAML1, CDH18, PCDH15, ROBO2	2.3E-2	5.8E-1
GO	integral component of membrane	ADAM12, ABCC1, AGFG1, DSCAML1, LRP1B, NDC1, VANGL1, ANO2, ARSF, CDH18, CERS3, CNTN6, CNTNAP5, DAGLB, DNAH10, ERMP1, GABRB1, GRID2, IL1RAPL1, LRRC4C, PTCH1, GALNT13, KCNMA1, POMGNT2, PTPRD, PTPRK, PCDH15, RTP4, ROBO2, RYR2, SELP, SORCS2, TACR1, TMCC1, TMEM132C, TMEM185A	4.5E-2	5.6E-1
GO	intracellular signal transduction	ARL15, AGAP1, DAB1, DAB2IP, ERC1, GARNL3, MECOM, RAB30, ARHGEF18, SSX2IP, ANK2, ARRB1, DIABLO, EDA, FOXO1, GUCY1A2 LRRC4C, PIK3CA, PLCB1, RIPK1, RYR2, SELP	1.1E-2	4.6E-1
GO	kidney development	EYA1, PTCH1, PCSK5, ROBO2, SULF1	4.7E-2	7.0E-1
GO	ligand-gated channel activity	GABRB1, GRID2, KCNJ3, RYR2	4.5E-2	7.0E-1
GO	ligand-gated ion channel activity	GABRB1, GRID2, KCNJ3, RYR2	4.5E-2	7.0E-1
GO	locomotion	DAB1, DAB2IP, SRGAP2, SSX2IP, DMRT1, DNAH11, PHACTR1, PLCB1, POMGNT2, PTPRK, ROBO2, SELP, SEMA6D, SULF1	3.1E-2	6.3E-1
GO	lung epithelium development	EYA1, FOXP2, THRB	2.3E-2	5.8E-1
GO	membrane part	ADAM12, ABCC1, AGFG1, DAB2IP, DSCAML1, FKBP9, LRP1B, NDC1, SRGAP2, VANGL1, ANK2, ANO2, ARSF, CDH18, CERS3, CNTN6,	8.0E-3	3.9E-1

		CNTNAP5, DIABLO, DAGLB, DLG2, DNAH10, EDA, ERMP1, EXOC4, GABRB1, GRID2, IL1RAPL1, LRRC4C, PTCH1, PIK3CA, GALNT13, KCNMA1, KCNJ3, POMGNT2, PTPRD, PTPRK, PCDH15, RIPK1, RTP4, ROBO2, RYR2, SELP, SEMA6D, SORCS2, SULF1, TACR1, TENM2, TMCC1, TMEM132C, TMEM185A		
GO	molecular function regulator	AGFG1, AGAP1, DAB2IP, GARNL3, ARHGAP20, ARHGAP6, ARHGEF18, SRGAP2, ARRB1, LRRC4C, PHACTR1, PHACTR2, PIK3CA, PLCB1	8.3E-3	3.8E-1
GO	movement of cell or subcellular component	DAB1, DAB2IP, SRGAP2, SSX2IP, ANK2, DMRT1, DNAH10, DNAH11, PHACTR1, PLCB1, POMGNT2, PTPRK, ROBO2, RYR2, SELP, SEMA6D, SULF1	8.2E-3	4.0E-1
GO	Multicellular organism development	AFF3, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SRGAP2, VANGL1, ANK2, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	6.4E-5	1.3E-1
GO	multicellular organismal process	AFF3, AGFG1, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, NDC1, PPARGC1A, SRGAP2, TNNI3K, VANGL1, ANK2, ARRB1, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DIAPH2, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PDE1C, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, RTP4, ROBO2, RYR2, SELP, SEMA6D, SOHLH2, SULF1, TACR1, TENM2, THRB	2.3E-4	7.1E-2
GO	muscle contraction	ANK2, KCNMA1, RYR2, SULF1, TACR1	4.2E-2	6.9E-1
GO	negative regulation of cell migration	DAB2IP, SRGAP2, PLCB1, PTPRK, SEMA6D, SULF1	4.9E-3	3.6E-1
GO	negative regulation of cell motility	DAB2IP, SRGAP2, PLCB1, PTPRK, SEMA6D, SULF1	5.9E-3	3.8E-1
GO	negative regulation of	DAB2, SRGAP2, PLCB1, PTPRK, SEMA6D, SULF1	1.1E-2	4.6E-1

	cellular component movement			
GO	negative regulation of epithelial cell proliferation	DAB2IP, CDK6, PTCH1, PTPRK, SULF1	3.0E-3	3.0E-1
GO	negative regulation of locomotion	DAB2IP, SRGAP2, PLCB1, PTPRK, ROBO2, SEMA6D, SULF1	2.2E-3	2.8E-1
GO	negative regulation of response to stimulus	DAB1, DAB2IP, EYA1, MECOM, ARRB1, CDK6, FOXO1, LRRC4C, PTCH1, RIPK1, ROBO2, SEMA6D, SULF1	3.1E-2	6.2E-1
GO	nervous system development	DAB1, DAB2IP, DSCAML1, EYA1, SRGAP2, CTNND2, CNTN6, CDK6, DAGLB, DLG2, FOXP2, LRRC4C, MAP2, PTCH1, PLCB1, KCNMA1, POMGNT2, PCDH15, ROBO2, SEMA6D, SULF1, TENM2	4.6E-4	1.2E-1
GO	neurogenesis	DAB1, DAB2IP, EYA1, SRGAP2, CTNND2, CNTN6, CDK6, DAGLB, LRRC4C, MAP2, PTCH1, KCNMA1, POMGNT2, PCDH15, ROBO2, SEMA6D, TENM2	1.3E-3	2.4E-1
GO	neuron development	DAB1, DAB2IP, SRGAP2, CTNND2, LRRC4C, MAP2, PCDH15, ROBO2, SEMA6D, TENM2	3.7E-2	6.7E-1
GO	neuron differentiation	DAB1, DAB2IP, EYA1, SRGAP2, CTNND2, CNTN6, LRRC4C, MAP2, PTCH1, KCNMA1, PCDH15, ROBO2, SEMA6D, TENM2	4.9E-3	3.5E-1
GO	neuron part	AGFG1, DAB2IP, ERC1, SRGAP2, CTNND2, DLG2, EXOC4, MAP2, KCNMA1, PTPRK, PCDH15, ROBO2, TENM2	3.2E-3	3.3E-1
GO	neuron projection	DAB2IP, DLG2, EXOC4, MAP2, KCNMA1, PTPRK, ROBO2, TENM2	3.8E-2	5.5E-1
GO	neuron projection morphogenesis	DAB1, DAB2IP, SRGAP2, CTNND2, MAP2, ROBO2, SEMA6D	4.6E-2	7.0E-1
GO	nucleoside-triphosphatase regulator activity	AGFG1, AGAP1, DAB2IP, GARNL3, ARHGAP20, ARHGAP6, SRGAP2, ARRB1, PLCB1	8.9E-5	1.0E-2
GO	pallium development	DAB1, DAB2IP, SRGAP2, FOXP2, PLCB1	7.1E-3	3.7E-1
GO	passive transmembrane transporter activity	ANO2, GABRB1, GRID2, MTMR6, KCNMA1, KCNJ3, RYR2	4.9E-2	7.1E-1

GO	plasma membrane	ADAM12, ABCC1, DAB2IP, NDC1, SRGAP2, VANGL1, ANK2, ANO2, CDH18, DIABLO, DAGLB, DLG2, EDA, EXOC4, GABRB1, GRID2, GUCY1A2, IL1RAPL1, NBEA, PARD6B, PTCH1, PIK3CA, KCNMA1, KCNJ3, PTPRK, PCDH15, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	1.8E-2	5.2E-1
GO	plasma membrane part	ABCC1, DAB2IP, SRGAP2, VANGL1, ANK2, DIABLO, DLG2, EDA, EXOC4, GRID2, PTCH1, KCNMA1, KCNJ3, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, TACR1, TENM2	6.8E-3	4.3E-1
GO	plasma membrane region	ABCC1, SRGAP2, ANK2, DLG2, GRID2, PTCH1, KCNMA1, ROBO2	4.2E-2	5.6E-1
GO	positive regulation of gene expression	DAB2IP, EYA1, MECOM, PPARGC1A, SMYD3, ZFP90, ANK2, ARRB1, CDK6, EDA, FOXO1, PTCH1, PLCB1, RIPK1, THRB	1.3E-2	4.8E-1
GO	positive regulation of nucleic acid-templated transcription	DAB2IP, EYA1, MECOM, PPARGC1A, SMYD3, ZFP90, ARRB1, FOXO1, PTCH1, PLCB1, RIPK1, THRB	3.0E-2	6.2E-1
GO	positive regulation of nucleobase-containing compound metabolic process	DAB2IP, EYA1, MECOM, PPARGC1A, SMYD3, ZFP90, ARRB1, FOXO1, GUCY1A2, PTCH1, PLCB1, RIPK1, THRB	4.6E-2	7.5E-1
GO	positive regulation of RNA biosynthetic process	ARHGAP6, ARRB1, PHACTR1, PCDH15	3.2E-2	6.3E-1
GO	positive regulation of RNA metabolic process	DAB2IP, EYA1, MECOM, PPARGC1A, SMYD3, ZFP90, ARRB1, FOXO1, PTCH1, PLCB1, RIPK1, THRB	4.0E-2	6.8E-1
GO	positive regulation of transcription, DNA-templated	DAB2IP, EYA1, MECOM, PPARGC1A, SMYD3, ZFP90, ARRB1, FOXO1, PTCH1, PLCB1, RIPK1, THRB	3.0E-2	6.2E-1
GO	postsynaptic membrane	SRGAP2, DLG2, GRID2, KCNMA1	1.2E-2	3.9E-1
GO	primary cilium	ANO2, DNAH11, PTCH1, PTPRK, PCDH15	1.3E-2	3.6E-1
GO	protein complex localization	NDC1, SSX2IP, DLG2, EDA	2.0E-2	5.6E-1

GO	protein tyrosine phosphatase activity	EYA1, MTMR6, PTPRD, PTPRK	2.1E-2	6.0E-1
GO	reflex	FOXP2, KCNMA1, PCDH15	9.3E-3	4.2E-1
GO	regulation of cardiac conduction	TNNI3K, ANK2, RYR2	3.6E-3	3.1E-1
GO	regulation of cellular response to stress	DAB2IP, EYA1, MECOM, CDK6, FOXO1, PLCB1, RIPK1	4.2E-2	6.9E-1
GO	regulation of epithelial cell proliferation	DAB2IP, EYA1, CDK6, FOXP2, PTCH1, PTPRK, SULF1	3.3E-3	3.1E-1
GO	regulation of heart contraction	TNNI3K, ANK2, RYR2, THRB	4.4E-2	7.0E-1
GO	regulation of locomotion	DAB2IP, SRGAP2, SSX2IP, PLCB1, PTPRK, ROBO2, SELP, SEMA6D, SULF1	3.8E-2	6.7E-1
GO	regulation of multicellular organismal process	DAB1, DAB2IP, EYA1, PPARGC1A, TNNI3K, ANK2, ARRB1, CDK6, DMRT1, FOXP2, LRRC4C, PTCH1, PIK3CA, PLCB1, KCNMA1, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, THRB	6.0E-3	3.8E-1
GO	regulation of muscle contraction	ANK2, KCNMA1, RYR2, TACR1	2.6E-2	6.1E-1
GO	regulation of muscle system process	ANK2, KCNMA1, RYR2, TACR1	4.4E-2	7.0E-1
GO	regulation of sequence-specific DNA binding transcription factor activity	DAB2IP, PPARGC1A, ARRB1, EDA, PTCH1, RIPK1	2.8E-2	6.2E-1
GO	regulation of small GTPase mediated signal transduction	DAB2IP, GARNL3, ARHGEF18, SSX2IP, ARRB1	3.8E-2	6.7E-1
GO	regulation of stress-activated MAPK cascade	DAB2IP, MECOM, FOXO1, PLCB1, RIPK1	1.9E-2	5.5E-1
GO	regulation of stress-activated protein kinase signaling cascade	DAB2IP, MECOM, FOXO1, PLCB1, RIPK1	1.9E-2	5.5E-1
GO	regulation of system process	TNNI3K, ANK2, KCNMA1, RYR2, TACR1, THRB	4.3E-2	6.9E-1
GO	regulation of Wnt signaling pathway	DAB2IP, CTNND2, EDA, FOXO1, SULF1	4.7E-2	7.0E-1

GO	signal transduction	ARL15, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, MECOM, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SRGAP2, SSX2IP, ANK2, ARRB1, CTNND2, CDK6, DIABLO, DMRT1, EDA, FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, PTCH1, PIK3CA, PDE1C, PLCB1, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	6.4E-3	3.7E-1
GO	signaling	ARL15, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, MECOM, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SRGAP2, SSX2IP, TNNI3K, ANK2, ARRB1, CTNND2, CDK6, DIABLO, DLG2, DMRT1, EDA, FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, PTCH1, PIK3CA, PDE1C, PLCB1, KCNMA1, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	4.2E-3	3.3E-1
GO	single organism signaling	ARL15, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, MECOM, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SRGAP2, SSX2IP, ANK2, ARRB1, CTNND2, CDK6, DIABLO, DLG2, DMRT1, EDA, FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, PTCH1, PIK3CA, PDE1C, PLCB1, KCNMA1, PTPRK, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SULF1, TACR1, TENM2	6.8E-3	3.7E-1
GO	single-multicellular organism process	AFF3, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SRGAP2, VANGL1, ANK2, ARRB1, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SELP, SEMA6D, SOHLH2, SULF1, TACR1, TENM2, THRB	7.0E-4	1.6E-1
GO	single-organism cellular process	PFKFB2, ARL15, ABCC1, AGFG1, AGAP1, DAB1, DAB2IP, ERC1, EYA1, GNA14, GARNL3, LARP1B, MECOM, NDC1, NLRP8, PPARGC1A, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SMYD3, SRGAP2, SSX2IP, ACOT12, ANK2, ARRB1, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DIABLO, DAGLB, DIAPH2, DLG2, DMRT1, DNAH10, DNAH11, EDA, EXOC4,	1.2E-3	2.4E-1

		FOXO1, GUCY1A2, IL1RAPL1, LRRC4C, MAP2, MYH13, MTMR6, PTCH1, PHACTR1, PIK3CA, PIP5K1B, PDE1C, PLCB1, GALNT13, KCNMA1, KCNJ3, PCSK5, POMGNT2, PTPRK, PCDH15, RIPK1, RTP4, ROBO2, RYR2, SCFD2, SELP, SEMA6D, SOHLH2, SULF1, TACR1, TENM2, THRB, USP37		
GO	single-organism developmental process	AFF3, AGFG1, DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SMYD3, SRGAP2, SSX2IP, VANGL1, ANK2, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	1.3E-4	9.3E-2
GO	single-organism process	PFKFB2, ARL15, AFF3, ABCC1, AGFG1, AGAP1, DAB1, DAB2IP, DSCAML1, ERC1, EYA1, GNA14, GARNL3, KIAA1217, LARP1B, MECOM, NDC1, NLRP8, PPARGC1A, RAB30, ARHGAP20, ARHGAP6, ARHGEF18, SMYD3, SRGAP2, SSX2IP, VANGL1, ACOT12, ANK2, ARRB1, CTNND2, CERS3, COL12A1, CNTN6, CDK6, DIABLO, DAGLB, DIAPH2, DLG2, DMRT1, DNAH10, DNAH11, EDA, EXOC4, FOXO1, FOXP2, GUCY1A2, IL1RAPL1, LRRC4C, MAP2, MYH13, MTMR6, PTCH1, PHACTR1, PIK3CA, PIP5K1B, PDE1C, PLCB1, GALNT13, KCNMA1, KCNJ3, PCSK5, POMGNT2, PTPRK, PCDH15, RIPK1, RTP4, ROBO2, RYR2, SCFD2, SELP, SEMA6D, SOHLH2, SULF1, TACR1, TENM2, THRB, USP37	1.4E-3	2.3E-1
GO	small GTPase mediated signal transduction	ARL15, AGAP1, DAB1, DAB2IP, RAB30, ARHGEF18, SSX2IP, ARRB1	2.4E-2	5.9E-1
GO	stress fiber assembly	ARHGAP6, ARRB1, PHACTR1	4.6E-2	7.0E-1
GO	stress-activated protein kinase signaling cascade	DAB2IP, MECOM, FOXO1, PLCB1, RIPK1	2.7E-2	6.1E-1
GO	synaptic membrane	SRGAP2, DLG2, GRID2, KCNMA1	2.6E-2	4.8E-1

GO	system development	DAB1, DAB2IP, DSCAML1, EYA1, KIAA1217, PPARGC1A, SRGAP2, ANK2, CTNND2, CERS3, CNTN6, CDK6, DAGLB, DLG2, DMRT1, DNAH11, EDA, EXOC4, FOXO1, FOXP2, LRRC4C, MAP2, PTCH1, PIK3CA, PLCB1, KCNMA1, PCSK5, POMGNT2, PCDH15, RIPK1, ROBO2, RYR2, SEMA6D, SOHLH2, SULF1, TENM2, THRB	1.3E-4	7.3E-2
GO	telencephalon cell migration	DAB1, DAB2IP, SRGAP2	5.0E-2	7.6E-1
GO	telencephalon development	DAB1, DAB2IP, SRGAP2, FOXP2, PLCB1	2.7E-2	6.1E-1
GO	telencephalon glial cell migration	DAB1, DAB2IP, SRGAP2	6.9E-3	3.7E-1
GO	tissue development	DAB2IP, EYA1, PPARGC1A, CERS3, COL12A1, CDK6, DMRT1, EDA, EXOC4, FOXP2, PTCH1, PIK3CA, KCNMA1, PCDH15, ROBO2, RYR2, SEMA6D, SULF1, THRB	2.0E-3	2.8E-1
GO	tube development	DAB2IP, EYA1, EDA, FOXP2, PTCH1, PCSK5, ROBO2, RYR2, THRB	1.3E-2	4.9E-1
GO	walking behavior	DAB1, KCNMA1, PCDH15	1.8E-2	5.4E-1
GO	Wnt signaling pathway	DAB2IP, CTNND2, EDA, FOXO1, RYR2, SULF1	3.5E-2	6.6E-1

Appendix VII.

References for each gene summary in Appendix IX.

Gene	Reference
ABCC1	(Beedholm-Ebsen et al., 2010)
ACOT12	(Horibata, Ando, Itoh, & Sugimoto, 2013)
ADAM12	(Kronqvist et al., 2002)
AFF3	(Feenstra et al., 2012)
AGAP1	(Luo et al., 2019)
AGFG1	(Yan, 2009)
ANK2	(Cunha et al., 2008)
ANO2	(Billig et al., 2011; Neureither et al., 2017)
ARHGAP20 (KIAA1391)	(Ferreira Pissarra, Olalla Saad, & Lazarini, 2018)
ARHGAP6	(Prakash et al., 2000)
ARHGEF18	(Arno et al., 2017)
ARL15	(Rocha et al., 2017)
ARMC3	(Pausch et al., 2016)
ARRB1	(Capurro et al., 2017)
ARSF	(Holmes, 2017)
BEGAIN	(Lamb et al., 2018)
BTAF1	(Yagi & Ip, 2005)
CDH18	(Bai et al., 2018)
CDK6	(Laurenti et al., 2015)
CELF2	(Ramalingam, 2012)
CERS3	(Uchida, 2014; Yokota, Masaki, Okano, & Tokudome, 2017; Zaki & Choate, 2018)
CFAP74	(Dong et al., 2017; McKenzie et al., 2015)
CNTN6	(Mercati et al., 2017)
CNTNAP5	(Ohye, Inagaki, Ozaki, Ikeda, & Kurahashi, 2014)
COL12A1	(Baker et al., 2017)(Punetha et al., 2017)
CTNNA2	(Ehlers et al., 2016; Fanjul-Fernández et al., 2013)
CTNND2	(Makunin et al., 2014)
DAB1	(Imai et al., 2017)
DAB2IP	(Valentino et al., 2017)
DAGLB	(Hsu et al., 2012)
DIABLO	(Verhagen et al., 2000)
DIAPH2	(Zhang et al., 2016)
DLG2	(Perrone et al., 2017)
DMRT1	(Zarkower et al., 2011; Zhao et al., 2015)
DNAH10	(Braathen et al., 2016)
DNAH11	(Lucas et al., 2012; Schwabe et al., 2008; Shin et al., 2017)
DSCAML1	(Fuerst et al., 2009)
EDA	(Tucker & Sharpe, 2004)
EDIL3	(Oh et al., 2017)
ERC1 (ELKS)	(Airik et al., 2016; Sigala, 2004)
ERMP1	(Cisternino et al., 2013)
EXOC4 (SEC8)	(Anitei et al., 2006)

EYA1	(Namba et al., 2001; Rebay, Silver, & Tootle, 2005; Zou et al., 2008)
EYS	(Hashmi et al., 2018)
FAM114A1	(Zhang et al., 2014)
FAM155A	(Lu et al., 2019)
FKBP9	(Ong et al., 2013)
Foxo1	(Munekata & Sakamoto, 2009)
FOXP2	(French et al., 2012)
GABRB1	(Halder et al., 2015)
GALNT13	(Desai et al., 2013)
GARNL3	(Lisowski et al., 2012)
GLT1D1	(Joo et al., 2018; Krzeminski et al., 2016)
GNA14	(Wang et al., 2018)
GPC6	(Capurro et al., 2017)
GRID2	(Selimi et al., 2003)
GTDC1	(Aksoy et al., 2017)
GUCY1A2	(Lightfoot, 2013)
IL1RAPL1	(Yasumura et al., 2015)
IST1	(Bajorek et al., 2009)
KCNJ3	(Yamada et al., 2019)
KCNMA1	(Bentzen et al., 2014)
KIAA0825	(Ullah et al., 2019)
KIAA1217	(Gershovich, Gershovich, & Buravkova, 2013; Karasugi et al., 2009)
LARP1B	(Ye et al., 2016)
LOC100687178	(Shimoyama et al., 2015)
LOC100688223 (FUCA1)	(Valero-Rubio et al., 2018)
LOC102153862	(Shimoyama et al., 2015)
LOC10251436	(Shimoyama et al., 2015)
LOC106559409	(Shimoyama et al., 2015)
LOC106559439	(Shimoyama et al., 2015)
LOC111093380	(Shimoyama et al., 2015)
LOC111094857	(Shimoyama et al., 2015)
LOC111098341	(Shimoyama et al., 2015)
LOC111098526	(Shimoyama et al., 2015)
LOC478449	(Shimoyama et al., 2015)
LOC487628 (THSD4)	(McDaniel, Li, Tordoff, Bachmanov, & Reed, 2006)
LOC487987	(Shimoyama et al., 2015)
LOC490812 (neurexin-3)	(Kelai et al., 2008)
LOC610994 (Cryzl2)	(Shimoyama et al., 2015)
LRP1B	(Liu, Musco, Lisitsina, Yaklichkin, & Lisitsyn, 2000)
LRRC4C	(Li et al., 2014)
MAP2	(Gumy et al., 2017)
MECOM	(Zhang et al., 2011)
MTMR6	(Harris, Parry, Westwick, & Ward, 2008)
MUC19	(Weitnauer et al., 2016)
MYH13	(Briggs & Schachat, 2002)
NBEA	(Nuytens et al., 2013),
NDC1	(Lai et al., 2016)
NKAIN2	(Mao et al., 2016)

NLRP8	(Chu et al., 2016)
NOL4	(Takayanagi-Kiya, Kiya, Kunieda, & Kubo, 2017)
NRG3	(Howard, 2008), (Zhang et al., 1997)
NTM	(McNamee, Reed, Howard, Lodge, & Moss, 2002)
PARD6B	(Alarcon, 2010)
PCDH15	(Ahmed et al., 2008; Murcia & Woychik, 2001; Schweizer, et al., 2016)
PCSK5	(Antenos et al., 2011)
PDE1C	(Wang et al., 2018)
PFKFB2	(Sreedhar, Petruska, Miriyala, Panchatcharam, & Zhao, 2017)
PHACTR1	(Wiezlak et al., 2012)
PHACTR2	(Kim et al., 2012)
PIK3CA (PI3K)	(Karakas, Bachman, & Park, 2006; Waardenberg et al., 2011)
PLCB1	(Vasco et al., 2012)
POMGNT2	(Nakagawa, Yagi, Kato, Takematsu, & Oka, 2015)
PPARGC1A	(Nakayama & Iwamoto, 2017)
PRKN	(Lockhart, O'Farrell, & Farrer, 2004)
PTCH1	(Lopez-Rios et al., 2014; Petit, Sears, & Ahituv, 2017)
PTPRD	(Choucair et al., 2015; Veeriah et al., 2009)
PTPRK	(Xu et al., 2015)
RAB30	(Oda et al., 2016; Thomas, Rousset, & Noselli, 2009)
RIPK1	(Peterson et al., 2017)
ROB2	(Shimoyama et al., 2015)
ROBO2	(Plump et al., 2002)
RTP4	(Gupta et al., 2010)
RYR2	(Diviani et al., 2013; Wenping Zhang et al., 2014)
SAMD12	(Cen et al., 2018)
SCFD2	(Ricketts et al., 2015)
SELP	(Volcik et al., 2009)
SEMA6D	(Kang & Kumanogoh, 2013)
SETD5	(Deliu et al., 2018)
SLX4IP (C20Orf94)	(Svendsen et al., 2009)
SMYD3	(Klinger et al., 2014)
SOHLH2	(Shin et al., 2017)
SorCS2	(Glerup et al., 2016; O'Rourke & Boeckx, 2018)
SRGAP2	(Tiwary, 2016)
SSX2IP	(Klinger et al., 2014)
SULF1	(Freeman et al., 2015)
TACR1 (NK1R)	(Douglas & Leeman, 2011) (Hoppe et al., 2018a)
TCERG1	(Montes et al., 2011)
TDRD3	(Yang et al., 2010)
TENM2	(Tews et al., 2017)
THRB	(Tigano, Reiertsen, Walters & Friesen 2018)
TLL2	(de Mooij-van Malsen et al., 2013)
TMCC1	(Zhang *et al.*, 2014)
TMEM132C	(Son et al., 2015)
TMEM185A	(Sheen et al., 2007)
TNNI3K	(Wang *et al.*, 2017)
TOX	(Chen et al., 2018)
TSHZ2	(Santos et al., 2010)
USP37	(Saxena & Kumar, 2014)

VANGL1	(DiTommaso et al., 2014; Kibar et al., 2009)
WDR41	(Hu et al., 2016)
ZFAND3	(Otake et al., 2011)
ZFP90	(Liu et al., 2018)

Appendix VIII.

Allele frequency from each of the 161 SNPs used in this study. Linked at allele freq 161snps K3 overlap Appendix VIII.xlsx (see excel).

Appendix IX. Candidate gene found in this study. A phenotypic trait shown to be influenced by each one of these genes by a previous study along with the system/organism used to gather this information. Finally, the FST for each SNP found in this study for each particular gene shown is also displayed. For more detail see Appendix III, IV, V, VI, VII, and VIII.

Gene	System/organism used or other (e.g. tissue, cell culture inferred)	Example of effect/phenotype	Category (involved in)	FST (One per SNP)
ABCC1 (MRP)	Mice	Transports cobalamin (Vitamin B12) into cells. Vital for ATP production.	Physiology	0.21365
ACOT12	Other	Involved in a variety of metabolism aspects in mammalian cells	Metabolism	0.2312
ADAM12	Mice	Involved with skeletal muscle development and maintenance	Development	0.2556
AFF3	Mice	Involved in dorsoventral patterning in embryonic limb development	Development	0.22475
AGAP1	Other	Plays a role in the cytoskeleton, cellular membrane trafficking and with the help of a kinesin (Kif2A) has also recently shown to affect actin dependent cell movement and actin remodelling	Physiology	0.20635
AGFG1	Mice	Vital in the process of acrosome biogenesis which is essential for healthy fertile sperm production	Reproduction	
ANK2	Humans, Mice	Involved with regulating the rhythm of cardiac muscle contraction	Physiology	0.23227, 0.27761
ANO2	Mice	Phototransduction, olfactory transduction, smooth muscle contact. Also involved in motor learning	Sensory/ Behaviour	0.34441
ARHGAP20	Other	Expressed predominantly in the brain where it plays a role in regulating the Rho-family GTPases and neurite outgrowth	Neuronal	0.18776
ARHGAP6	Mice, Other	Plays a role in regulating cell morphology and the cytoskeleton, also probably plays a role in neurodevelopmental	Physiology / Development	0.19711

ARHGEF18	Medaka fish, Human	Involved with retinal development and maintenance of photoreceptor viability	Senses/ Development	0.25122
ARL15	Mice	Involved with adipocyte differentiation and adiponectin secretion.	Metabolism	0.28038
ARMC3	Bovine	Spermatogenesis and infertility	Reproduction	0.28726
ARRB1	Vertebrates, Other	Involved in phototransduction in rod cells	Physiology/ Sensory	0.19891
ARSF	Other	Little known, thought to play a role in lysosomal sphingolipid metabolism	Metabolism	0.22894
BEGAIN	Sheep	Genomic imprinting (via epigenetic effects) in the germ line cells (hypermethylated in mature sperm, complete methylation only in late stages of oocyte development in sheep)	Reproduction	0.24916
BTAF1	Other	Is an ATPases/helicases that interact with TATA binding protein (TBP). Involved in activating or suppressing transcription	Metabolism	0.20129
CDH18	Other	Involved in neuronal biological processes. Also, tumour suppressor.	Nervous system (Tumour suppressor)	0.26389, 0.24355, 0.31583
CDK6	Mice	Involved in cell cycle e.g Regulates Hematopoietic Stem Cells	Physiology	0.20758
CELF2	Human	mRNA splicer Tumour suppressor also involved with heart development, nervous system, and striated muscle.	Tumour suppressor	0.25968
CERS3	Mice, Other	Involved in lipid metabolism and synthesis of keratinocytes (barrier against pathogens). Also, possibly involved with sperm synthesis	Metabolism (reproduction)	
CFAP74	Mice	Involved with tissues that use motile cells including testes, lung and brain.	Possibly Reproduction/ Senses (olfaction)	0.22595
CNTN6	Mice	involved in sensory-motor neuronal pathways (coordination possible hearing loss)	Development/ Senses	0.15793
CNTNAP5	Other	Belongs to the neurexin superfamily with unknown function, however CNTNAP5 is a brain-specific gene	Unknown	0.2024

COL12A1 (EX)	Other, Humans, Dogs	Involved with collagen. Associated with ligaments ruptures and muscle weakness	Development	0.19794
CTNNA2	Mice	Involved with behavioural responses. Also, Tumour suppressor	Behaviour (Tumour suppressor)	0.32874
CTNND2	Fox	Protooncogene	Protooncogene	0.29094
DAB1	Mice	Involved in locomotor activity, working, spatial, and contextual fear memory, cortical development and brain functions	Behaviour/ neuronal	0.19683
DAB2IP	Mice	Many cancers	Tumour suppressor	0.23418
DAGLB	Other	Involved in macrophage proinflammatory responses	Immune	0.24098, 0.17936, 0.30031
DIABLO	Other	Involved with apoptosis	Physiology	0.2426
DIAPH2	Human	Involved with development of ovary	Reproduction	0.24382
DLG2	Mice, Humans (Highly conserved in vertebrates – 100MY), Other	Involved in complex learning, cognitive flexibility and attention. Various cognitive learning and attention factors especially associative learning.	Behaviour	0.21976
DMRT1	Mice, Chicken, Fish (vertebrates-conserved), Other	Critical for formation of male gonads (subordinate to SRY gene), production and survival of sperm. (May also be a tumour suppressor)	Reproduction/ (tumour suppressor)	0.28436
DNAH10	Other	force-generating protein of respiratory cilia	Physiology	0.25021
DNAH11	Other	Essential for motility of cilia, lack of can cause; Asthenospermia up to Infertility (not sterility) Lack of movement of respiratory cilia and other related primary ciliary dyskinesia diseases	Reproduction/ Health	0.18863
DSCAML1	Other	Involved with rental development	Development	0.23345
EDA	Mice	Development of teeth	Development	0.26037
EDIL3	Mice	Involved with embryonic cranial development and osteoblast differentiation	Development	0.22222

ERC1 (ELKS)	Mice, Other	A component in the NF-κB signalling cascade which is involved with expression in many cell types during fundamental biological processes, including the immune response, apoptosis, oncogenesis, and development	Physiology/ Development	0.18518
ERMP1	Rats	Organization of somatic cells and oocytes into discrete follicular structures	Reproduction	0.25108
EXOC4 (SEC8)	Rats, Other	Plays a role in the nervous system development	Development	0.29073
EYA1	Humans, Mice	Ear development and hearing	Development / Senses	0.24657
EYS	Human	Visual function	Sensory	0.19057, 0.20974, 0.23434
FAM114A1	Other	Involved with early neuronal development	Development /Neuronal	0.16931
FAM155A	Other	Play a role in tooth development	Development	0.26484
FKBP9	Other	Possibly a chaperone (folding proteins)	Physiology	0.20583
Foxo1	Mice	Regulates adipocyte differentiation	Metabolism (Thermogenesis)	0.2238
FOXP2	Other, Mice	Neural plasticity in cortico-basal ganglia circuits underlying the sensory-guided motor learning in animal models	Behaviour	0.18123
GABRB1	Other	Involved with early brain development	Neural/ Development	0.27116
GALNT13	Other	Regulated by hypoxia, may be involved with pulmonary vascular remodelling	Physiology/ Metabolism	0.25038
GARNL3	Mice	Involved with regulation of signal transduction and GTPase activity, upregulated in swim stress-induced analgesia	Physiology/ Metabolism	0.17714
GLT1D1	Human	Candidate oncogene of colorectal cancer	Oncogene (candidate)	0.27705
GNA14	Mice	Oncogene	Oncogene	0.25374
GPC6	Mice	Involved in bone mineral content, bone formation and skeletal signalling	Development	0.17357
GRID2	Mice	Plays an important role in modulating neuronal death, neurodevelopmental, neurophysiological, and neuropathological processes	Development/ Behaviour	0.24384

GTDC1	Zebrafish, Human, Other	Involved in the development of the nervous system (neural development)	Development	0.20277
GUCY1A2	Mice	Involved with regulating activity, see paper for hypothesis of NS	Metabolism/ Behaviour	0.16394
IL1RAPL1	Mice	Involved in synaptic plasticity, brain functions including learning, memory, behavioural flexibility, locomotor activity and anxiety.	Neuronal/ behavioural	0.2975
IST1	Other	Plays a role in cytokinesis	Physiology	0.1825
KCNJ3	Xenopus laevis, Human, Other	Involved in the regulation of cardiac muscle.	Physiology	0.317
KCNMA1	Mice, Rats, Guinea pigs	Involved with hearing frequencies	Sensory	0.19588
KIAA0825	Mice	Little known, mouse ortholog involved with limb development	Development	0.21816
KIAA1217	Other, Mice	Involved in the differentiation and development of the skeletal system especially the intervertebral discs (IVDs) in humans and mice	Development	0.20747
LARP1B	Other	Involved with transcription and/or mRNA translation (including post-transcription)	Metabolism/ Development	0.26771
LOC100688223 (FUCA1)	Other	Involved in immune response and keratinocyte differentiation/epidermal development	Immune/ Development	0.24955
LOC487628 (THSD4)	Other	Regulates adipogenesis	Metabolism	0.27576
LOC490812 (neurexin-3)	Other, Mice	A membrane protein that is involved with synaptic transmission, neurotransmitter secretion and cell adhesion	Neuronal/ Development	0.24683
LOC610994 (Cryzl2)	Other	Involved in oxidation-reduction processes	Metabolism	0.2073, 0.24593, 0.23012, 0.23341
LRP1B	Mice, Other	Many cancers especially lung cancer	Tumour suppressor	0.2454
LRRC4C	Mice, Humans	Tumour suppressor mainly expressed in the brain.	Tumour suppressor	0.23014, 0.20234, 0.22987
MAP2	Other	Directs axon cargo transport such as organelles or molecules along microtubules in nerve cell axons in sensory neurons	Physiology /Neuronal	0.16799

MECOM	Other	Hematopoietic stem cell (HSC) regulation. Also bone development	Physiology/ Development	0.22282
MTMR6	Other	May inhibit aspects of immune system such as T-cell activation	Immune	0.3484
MUC19	Mice, Human?, Canine	Gel-forming mucin that lubricates saliva and plays a role in reducing adherence and increasing clearance of bacteria which may be important in systemic health and other diseases.	Health (pathogens)	0.33765
MYH13	Other	Plays a role in ocular motility	Physiology/ Sensory	0.23215
NBEA	Mice	Involved in neurotransmitter release, synaptic functioning behavioural aspects	Behaviour	0.30959
NDC1	Mice	Involved with spermatogenesis	Development/ Reproduction	0.24954
NKAIN2	Other	Involved in prostate cancer	Tumour suppressor	0.15053
NLRP8	Some mammals, Other	Has been shown to be upregulated in *T. gondii* infections so may have an immunity role too.	Immune system	0.26844
NOL4	Mice, Honeybees, Drosophila	Plays a role in transactivating (upregulating gene expression) of Mlr1 and Mlr2 which is involved with the neural development.	Development	0.3107
NRG3 (ex)	Other	Promotes mammary differentiation through signalling. Also involved with developing nervous system	Development	0.24792
NTM	Rats	Involved in Regulating neurite growth of neurons	Development / Neuronal	0.21787
PARD6B	Mice	essential for normal blastocyst formation	Development	0.29171
PCDH15	(Vertebrate hair cells) Mice, Other	Involved with hearing, movement/balance. Also, vision	Sensory	0.31599
PCSK5	Mice	Regulates the maturation of the ovarian follicle (folliculogenesis). (also, early development)	Reproduction/ (Development)	0.27898
PDE1C	Other	By catalysing and hydrolysing intracellular second messengers, cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) regulation of homeostatic processes occurs such as	Sense/ Physiology	0.18529

		Ca2+ homeostasis which is vital for hearing		
PFKFB2	Mice, Other	Regulates glycolysis	Metabolism	0.20516
PHACTR1	Other	Important in assembling muscle tissue components actomyosin (actin and myosin). Also associated with melanoma progression, early-onset myocardial infarct, and endothelial cell tube formation.	Physiology	0.23914
PHACTR2	Mice, Other	Expressed in brain region which involves non-spatial associative learning through combining external sensory inputs with self-motivation stimuli	Neural	0.32701
PIK3CA (PI3K)	Mice, Rats	Vital regulator of the expression of genes essential for cardiac structure and Z-disc signalling. Also protects the heart in high stress and is critical for Z-disc alignment. Also, an oncogene.	Physiology/ Development/ Oncogene	0.23558
PIP5K1B	Other	Involved in neurite formation and by regulating a lipid secondary messenger (phosphoinositide PtdIns(4,5) P2) controls responses to cellular stresses	Metabolism	0.22925, 0.22711
PLCB1	Mice	Involved with post-natal cortical development, neural plasticity, Spatial memory, motor/sensorimotor gating and behavioural traits	Behaviour/ Neuronal	0.161
POMGNT2	Mice	Involved in early neuronal development	Development	0.30183
PPARGC1A	Mice, Human	Development of brown/beige adipocytes in response to cold stress leading to adaptive thermogenesis.	Metabolism (Thermogenesis)	0.25285
PRKN	Mice	Involved with the nervous systems and behaviour. A knock out mouse shows Behavioural deficits such as reduced exploratory behaviour	Behaviour	0.24366
PTCH1	Mice, Bovine	Involved in digit regulation and autopod symmetry	Development	0.19894
PTPRD	Mice	Involved in memory and learning. Also, involved in hearing and a tumour suppressor	Neuronal/ behaviour and Senses Also, Tumour suppressor	0.3546

PTPRK	Mice	Involved with regulating keratinocyte proliferation	Physiology	0.26509
RAB30	Drosophila	A GTPase that plays a role in morphogenesis. Also, may play a role in regulating autophagy in an immunological response	Development/ Immune	0.31107
RIPK1	Other	Plays a critical role in activating the innate immune system during bacterial infection.	Immune	0.28594
ROBO2	Zebrafish	Involved with retinal axon guidance	Development/ Senses	0.15934
RTP4	Mice	Upregulates as a defence to virus infection	Immune	0.30426
RYR2	Other	Increases Ca2 + mobilization from intracellular stores and promotes cardiac contraction. Associated with high altitude adaptation in humans and wolves	Physiology	0.18771
SAMD12 (EX)	Mice	Little known, possible immunological aspects. Loss of function may lead to neurotoxicity.	Immune ?	0.31564
SCFD2 (EX)	Dog	Little known, may be involved with protein transport in the eye	Physiology/ Senses	0.2932
SELP	Other	Involved with platelet adhesion, Leukocyte–endothelium and leukocyte–platelet interactions and immune (inflammation) response	Immune	0.17376
SEMA6D	Mice	Bone morphogenetic protein, and involved with heart development,	Development	0.1887
SETD5	Other	Regulates gene transcription of genes involved with hippocampal long-term memory, novel-object location memory, fear memory retention, long-term potentiation, impaired spatial learning, and adaptive learning Also, involved in mammalian embryonic development.	Development/ Behaviour	0.2066
SLX4IP (C20Orf94)	Other	May be involved with telomere homeostasis	Physiology	0.23461, 0.19837
SMYD3	Other	Involved with regulating transcription in oncogenes, homeobox genes and genes associated with cell-cycle	Physiology	0.20288

		regulation by forming a complex with RNA polymerase II		
SOHLH2	Mice	Regulates spermatogonial differentiation and is also essential for spermatogonial development. Also, essential regulator of oocyte differentiation	Reproduction	0.29178
SorCS2	Mice	Impaired formation of long-term memory, increased risk taking and stimulus seeking behaviour, enhanced susceptibility to stress and impaired prepulse inhibition.	Behaviour	0.18324, 0.20502
SRGAP2	Mammals	It has been previously shown to play an important in neuronal migration and differentiation which leads to the development of the brain cortex	Neuronal/ Development	0.24191
SSX2IP	Other	Essential role in normal functioning of primary cilia.	Physiology	0.2513
SULF1	Mice, Other	Involved with inner ear development/morphogenesis	Development	0.2379
TACR1 (NK1R)	Mice, Rats, Human	Many functions. Interacts with Substance-P receptor to activate nuclear factor-kappa-b (NF-κb) and activate and modulate proinflammatory cytokines and signals which are important in bacterial, viral, fungal, and parasitic diseases, as well as in immune system function. Also modulate stress and anxiety-related behaviour, fear, and aggression	Health (pathogens)/ Neuronal behaviour	0.23006
TCERG1	Other	Involved in transcriptional elongation and pre-mRNA splicing e.g. Regulates alternative splicing in an apoptotic gene Bcl-x	Metabolism	0.23765
TDRD3	Other	Recognises and reads histone methylarginine marks and is a transcriptional coactivator (increases gene expression)	Metabolism	0.18287
TENM2	Human	Regulates lipids. Maintains white adipose tissue, blocking the gene increases brown adipose tissue. Also involved with regulating synaptic connections	Metabolism (Possible Thermogenesis) Neuronal	0.24277

THRB	Mice	Encodes for receptors of which are important to lipid metabolism and adaptive thermogenesis.	Metabolism (Thermogenesis)	0.2501
TLL2	Mice, Humans	Avoidance behaviour in mice, bipolar disorder in humans. Also bone development/disease	Behaviour/ Development	0.20846
TMCC1	Other	Involved with endoplasmic reticulum (ER) organization	Physiology	0.21698
TMEM132C	Other, Human	Involved with the respiratory systems pulmonary function	Physiology	0.17396
TMEM185A	Other, Human	Unknown Phenotype	Unknown	0.23129
TNNI3K	Mice	Involved with cardiomyogenesis	Physiology/ Metabolism	0.24687
TOX	Other	Involved with regulating thrombocytes and T-cells	Immune	0.22547
TSHZ2	Zebrafish	Involved with hindbrain and neural retina development.	Development	0.2298
USP37	Other	Regulates the cell cycle at the G1/s boundary, and S and G2 phases through interactions with the APC/CDH1 complex and Cyclin A	Physiology	0.24618
VANGL1	Mice	Involved in pigmentation and correct formation of neural tube	Physiology/ Development	0.27671
WDR41	Mice	Involved with autophagy/lysosomal regulation	Immune system/ pathogens	0.24745
ZFAND3	Mice, tilapia (Fish)	Essential for spermatogenesis in mice. Possibly also sex determination, differentiation, and involved in male germ cell maturation in tilapia	Reproduction	0.27796
ZFP90	Mice	Regulates hematopoietic stem cells (HSCs) self-renewal and differentiation.	Physiology/ Metabolism (Development)	0.47391, 0.47986